A

**Project Report**

on

# YOLOv8 vs YOLOv9: Real-Time Object Detection in Live Football Broadcasts

submitted as partial fulfillment for the award of

# BACHELOR OF TECHNOLOGY DEGREE

SESSION 2024-25

in

## Computer Science and Engineering

By

Eva Mittal (2100290120077)

Aditya Agnihotri (2100290100012)

Devyanshi Srivastava (2100290100054)

**Under the supervision of**

Mr. Gaurav Parashar

## KIET Group of Institutions, Ghaziabad

Affiliated to

**Dr. A.P.J. Abdul Kalam Technical University, Lucknow**

(Formerly UPTU)

**May, 2025**

# DECLARATION

We hereby declare that this submission is our own work and that, to the best of our knowledge and belief, it contains no material previously published or written by another person nor material which to a substantial extent has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Signature
Name: Eva Mittal
Roll No.: 2100290120077
Date:

Signature
Name: Aditya Agnihotri
Roll No.: 2100290100012
Date:

Signature
Name: Devyanshi Srivastava
Roll No.: 2100290100054
Date:

# CERTIFICATE

This is to certify that Project Report entitled "YOLOv8 vs YOLOv9: Real-Time Object Detection in Live Football Broadcasts" which is submitted by Eva Mittal, Aditya Agnihotri, Devyanshi Srivastava in partial fulfillment of the requirement for the award of degree B. Tech. in Department of Computer Science & Engineering of Dr. A.P.J. Abdul Kalam Technical University, Lucknow is a record of the candidates own work carried out by them under my supervision. The matter embodied in this report is original and has not been submitted for the award of any other degree.

| | |
|---|---|
| **Mr. Gaurav Parashar** | **Dr. Vineet Sharma** |
| **(Associate Professor)** | **(Dean, CSE Department)** |

**Date:**

ACKNOWLEDGEMENT

It gives us a great sense of pleasure to present the report of the B. Tech Project undertaken during B. Tech. Final Year. We owe special debt of gratitude to Mr. Gaurav Parashar, Department of Computer Science & Engineering, KIET, Ghaziabad, for his constant support and guidance throughout the course of our work. His sincerity, thoroughness and perseverance have been a constant source of inspiration for us. It is only his cognizant efforts that our endeavors have seen light of the day.

We also take the opportunity to acknowledge the contribution of Dr. Vineet Sharma, Dean of the Department of Computer Science & Engineering, KIET, Ghaziabad, for his full support and assistance during the development of the project. We also do not like to miss the opportunity to acknowledge the contribution of all the faculty members of the department for their kind assistance and cooperation during the development of our project.

We also do not like to miss the opportunity to acknowledge the contribution of all faculty members, especially faculty/industry person/any person, of the department for their kind assistance and cooperation during the development of our project. Last but not the least, we acknowledge our friends for their contribution in the completion of the project.

Signature
Name: Eva Mittal
Roll No.: 2100290120077
Date:

Signature
Name: Devyanshi Srivastava
Roll No.: 2100290100054
Date:

Signature
Name: Aditya Agnihotri
Roll No.: 2100290100012
Date:

# ABSTRACT

Object detection has become a vital component of computer vision, with diverse applications ranging from surveillance and self-driving cars to sports analysis. In the sports arena, especially in football, real-time object detection systems play a crucial role in tracking players, observing ball movements, and improving tactical evaluations. This report offers a comparative analysis of two leading object detection models—YOLOv8 and YOLOv9—aimed at determining which model is more efficient and accurate for football-related object detection tasks.

The YOLO (You Only Look Once) family of algorithms is well-regarded for balancing detection speed with high accuracy. YOLOv8, developed by Ultralytics, includes improvements such as anchor-free detection, a decoupled head design, and adaptability for various tasks including classification, detection, and segmentation. On the other hand, YOLOv9 is the latest release in the YOLO series, introducing cutting-edge architectural components like the Generalized Efficient Layer Aggregation Network (GELAN) and Programmable Gradient Information (PGI) to enhance both accuracy and computational efficiency even more.

The research focuses on creating a tailored dataset derived from actual football games, showcasing a range of intricate situations such as player overlaps, motion blur, unusual lighting conditions, and rapidly moving objects. Both models underwent training on an identical dataset with the same hyperparameters to ensure a fair comparison. Performance evaluation was conducted using key metrics, that included mean Average Precision (mAP), Precision, Recall, F1-score, and Frames Per Second (FPS).

The findings indicate that YOLOv8 strikes a great balance between speed and precision, while YOLOv9 offers higher accuracy (especially noted by mAP@0.5 and mAP@0.5:0.95) and achieves quicker convergence in the training phase. Nevertheless, YOLOv8 displays somewhat enhanced performance in real-time inference on resource-constrained systems, which could be advantageous for deployment in edge circumstances.

This report contributes to the growing literature on deep learning for sports video analytics by evaluating recent advancements in object detection using a football-specific dataset. It

focuses on assessing the performance of modern models, such as YOLOv8 and YOLOv9, in handling the fast-paced and complex scenarios typical of football matches. The insights gained from this evaluation are valuable for researchers and developers aiming to build real-time, high-accuracy detection systems for applications like player tracking, tactical analysis, and smart broadcasting in dynamic sports environments.

# TABLE OF CONTENTS


| | Page no. |
|---|---|
| DECLARATION………………………………………………………….. | ii |
| CERTIFICATE……………………………………………………..….. | iii |
| ACKNOWLEDGEMENT…………………………………………......... | iv |
| ABSTRACT………………………………………………………….. | v |
| LIST OF FIGURES……………………………………………………... | x |
| LIST OF TABLES……………………………………………………… | xi |
| LIST OF ABBREVIATION……………………………………........... | xii |
| **CHAPTER 1 (INTRODUCTION)**……………………………………… | 1-7 |
| 1.1 Introduction……………………………………………………..….. | 1 |
| 1.2 Background and Motivation…………………………………………… | 1 |
| 1.3 Problem Statement………………………………………………….. | 2 |
| 1.4 Objectives of the Study………………………………………………... | 3 |
| 1.5 Scope of the Project…………………………………………………… | 4 |
| 1.6 Significance of the Project…………………………………………….. | 5 |
| 1.7 Structure of the Report………………………………………………… | 7 |
| **CHAPTER 2 (LITERATURE REVIEW)**…………………………….. | 8-10 |
| 2.1 Introduction………………………………………………………… | 8 |
| 2.2 Object Detection: Traditional to Deep Learning-Based Approaches….. | 8 |
| 2.3 Emergence and Evolution of YOLO Architecture……………………... | 8 |
| 2.4 YOLOv8 and Its Technical Innovations……………………………….. | 9 |
| 2.5 YOLOv9: Next-Generation Efficiency………………………………… | 9 |
| 2.6 YOLO in Sports Analytics……………………………………………… | 9 |
| 2.7 Summary of Gaps in Literature………………………………………… | 10 |

**CHAPTER 3 (PROPOSED METHODOLOGY)**……………………... 11-27
3.1 Overview………………………………………………………………. 11
3.2 System Architecture
3.2.1 Overview of the System Architecture……………………...... 11
3.2.2 Modules of the Architecture…………………………………. 12
3.3 Dataset Preparation
3.3.1 Data Collection……………………………………………..... 13
3.3.2 Data Annotation……………………………………………… 14
3.3.3 Data Augmentation Techniques……………………………… 14
3.4 Model Description
3.4.1 YOLOv8……………………………………………………… 15
3.4.2 YOLOv9……………………………………………………… 16
3.5 Training Procedure
3.5.1 Hardware and Software…………………………………...... 18
3.5.2 Hyperparameters…………………………………………….. 19
3.6 Evaluation Metrics…………………………………………………… 19
3.7 Loss Functions
3.7.1 YOLOv8 Loss Function…………………………………..... 20
3.7.2 YOLOv9 Loss Function…………………………………..... 22
3.7.3 Image Analysis for the loss function…………………………. 23
3.7.4 Loss Curves and Convergence Analysis……………………… 24
3.7.5 Comparative Summary……………………………………….. 24
3.8 Workflow Summary…………………………………………………. 25
3.9 Model Configuration Comparison……………………………….…..... 26

**CHAPTER 4 (RESULTS AND DISCUSSION)**……………………… 28-37
4.1 Introduction………………………………………………………. 28
4.2 Experimental Setup
4.2.1 Dataset Description…………………………………………... 28
4.2.2 Hardware and Software……………………………………… 29
4.3 Evaluation Metrics…………………………………………………... 29
4.4 Quantitative Results…………………………………………………. 31

4.5 Visual Results……………………………………………………………….. 31
4.6 Comparative Analysis………………………………………………………….. 35
4.7 Error Analysis………………………………………………………………….. 36
4.8 Practical Implications………………………………………………………….... 37
4.9 Summary………………………………………………………………………….. 37

**CHAPTER 5 (CONCLUSION AND FUTURE SCOPE)**…………….. 38-40
5.1 Conclusion………………………………………………………………... 38
5.2 Future Scope…………………………………………………………………. 39
5.3 Final Remarks……………………………………………………………….. 40

REFERENCES…………………………………………………………….. 41-43
APPENDIX………………………………………………………………… 44-58

# LIST OF FIGURES


| Figure No. | Description | Page No. |
|---|---|---|
| 3.1 | System Architecture | 11 |
| 3.2 | Dataset images | 13 |
| 3.3 | Labelled Images | 14 |
| 3.4 | YOLO v8 Architecture | 16 |
| 3.5 | YOLO v9 Architecture | 17 |
| 3.6 | Loss function for YOLO v8 | 21 |
| 3.7 | Loss function for YOLO v9 | 23 |
| 4.1 | Comparison of confusion matrices for YOLOv8 and YOLOv9 | 32 |
| 4.2 | Comparison of PR Curve for YOLOv8 and YOLOv9 | 33 |
| 4.3 | Comparison of F1 Curve for YOLOv8 and YOLOv9 | 34 |
| 4.4 | Comparison of Labels for YOLOv8 and YOLOv9 | 35 |

# LIST OF TABLES


| Figure No. | Description | Page No. |
|---|---|---|
| 3.1 | H/w and S/w specification | 18 |
| 3.2 | Specification for Hyperparameters | 19 |
| 3.3 | Loss function images description | 23 |
| 3.4 | Comparative Summary | 24 |
| 3.5 | Workflow Summary | 25 |
| 3.6 | Model Configuration Comaprison | 26 |
| 4.1 | Comparison of YOLO v8 and v9 results | 31 |

# LIST OF ABBREVIATIONS

| Abbreviation | Full Form |
|---|---|
| AI | Artificial Intelligence |
| API | Application Programming Interface |
| BCE | Binary Cross Entropy |
| CIoU | Complete Intersection over Union |
| CNN | Convolutional Neural Network |
| COCO | Common Objects in Context (Dataset) |
| CSP | Cross Stage Partial Network |
| DFL | Distribution Focal Loss |
| FPS | Frames Per Second |
| GIoU | Generalized Intersection over Union |
| IoU | Intersection over Union |
| mAP | Mean Average Precision |
| ML | Machine Learning |
| NMS | Non-Maximum Suppression |
| ONNX | Open Neural Network Exchange |
| PGI | Programmable Gradient Information |
| PyTorch | Python-based Machine Learning Framework |
| RepOptimizer | Re-parameterized Optimizer |
| TAL | Task-Aligned Learning |
| YOLO | You Only Look Once |
| YOLOv8 | YOLO version 8 |
| YOLOv9 | YOLO version 9 |

# CHAPTER 1

# INTRODUCTION

## 1.1 Introduction

In the fast-changing domain of computer vision, object detection has become a fundamental technology applicable to a wide array of real-world uses—from surveillance systems and self-driving cars to intelligent retail and sports analytics. Object detection performs both the identification and categorization of objects within an image, merging the challenges of classification and localization into a single task. Recently, the advancement of deep learning models, especially convolutional neural networks (CNNs), has greatly enhanced the effectiveness of object detection.

Among the many object detection algorithms developed, the "You Only Look Once" (YOLO) series has stood out for its exceptional speed and accuracy. The YOLO architecture has been in continuous development since being introduced, evolving from YOLOv1 to the most recent models, YOLOv8 and YOLOv9. Each iteration introduces new and exciting changes to the model in order to improve accuracy while reducing computational cost and allowing object detection to be more feasible and immediate.

This project centers around a comparative analysis of YOLOv8 and YOLOv9 with respect to object detection in images of football images. Reflecting the dynamic and fast nature of football games, effective and precise object detection is essential for automation systems to enable the analysis of games for player tracking, team analysis, and tactical evaluations.

## 1.2 Background and Motivation

Object detection in sports, especially in football, presents unique challenges such as high-speed object movement, frequent occlusions, varying lighting conditions, and crowded scenes. Conventional object detection models often struggle with such complexities. The YOLO family of models addresses these challenges by offering high-speed detection with increasing accuracy across newer versions.

YOLOv8, released by Ultralytics, brought significant architectural upgrades, including decoupled heads for classification and localization, better data augmentation techniques, and simplified model scaling. YOLOv9, the latest addition, further pushes the boundary by introducing new modules like the Generalized Efficient Layer Aggregation Network (GELAN) and an improved detection head known as DWConv SimOTA, making it even more suitable for real-time applications.

The motivation behind this project lies in the need to evaluate whether YOLOv9's improvements translate into measurable gains over YOLOv8 in a real-world use case like sports analytics, where detection precision and speed are paramount.

## 1.3 Problem Statement

Despite rapid advancements in deep learning, **real-time object detection in complex sports environments like football remains a significant challenge**. The fast-paced nature of the game, unpredictable player movements, occlusion, dynamic lighting, and camera shifts create difficult conditions for precise object recognition and tracking.

Although models in the **YOLO (You Only Look Once)** family are widely celebrated for their speed and accuracy, newer versions such as **YOLOv8** and **YOLOv9** are still under active research and adoption. These models claim to offer improved performance via architectural enhancements and training efficiency. However, **comparative performance evaluations specific to football analytics** — where latency and accuracy are both critical — are scarce.

Therefore, the problem this study addresses is:

**"How do YOLOv8 and YOLOv9 compare in real-time object detection tasks specific to football in terms of accuracy, inference speed, and computational cost?"**

This comparison aims to guide developers, researchers, and sports technologists in selecting the most suitable model for building intelligent and scalable sports analytics systems.

## 1.4 Objectives of the Study

The primary objective of this project is to **evaluate and compare the effectiveness of YOLOv8 and YOLOv9 models** for object detection tasks in the context of football. The specific goals are outlined below:

### 1.4.1 Model Training and Evaluation

- Train YOLOv8 and YOLOv9 on a football-specific dataset with objects such as players, footballs, and goalposts.
- Evaluate models using standard performance metrics: **mAP (mean Average Precision)**, **precision**, **recall**, and **F1-score**.

### 1.4.2 Comparative Analysis

- Compare both models on the basis of:
  - Detection accuracy and robustness under challenging visual conditions.
  - Real-time inference speed (FPS – frames per second).
  - Training time and model convergence.
  - Computational cost (CPU/GPU usage, RAM).

### 1.4.3 Dataset Preparation

- Prepare and annotate a custom dataset or leverage an existing football-specific dataset tailored for object detection.
- Apply preprocessing and augmentation techniques to improve model generalization.

### 1.4.4 Deployment Consideration

- Assess the deployability of both models in real-time or near-real-time football analytics systems.
- Discuss practical scenarios and use-cases for sports coaches, analysts, and referees.

### 1.4.5 Documentation and Reporting

- Provide reproducible results and detailed documentation for further academic or commercial development.

## 1.5 Scope of the Project

The scope of this project encompasses the **design, implementation, and comparative evaluation of object detection models** specifically in the context of football. While the primary focus lies in evaluating YOLOv8 and YOLOv9, the scope also involves a broader understanding of model behavior, dataset handling, and performance metrics under dynamic conditions. The boundaries and key inclusions are as follows:

### 1.5.1 Inclusions

- Focused evaluation of **YOLOv8 and YOLOv9** using a **football-specific dataset**.
- Consideration of both training and testing phases for a fair and thorough comparison.
- Implementation of standard performance metrics such as **mAP, precision, recall, F1-score, FPS**, and **IoU (Intersection over Union)**.
- Use of **custom-trained models** with domain-specific tuning for accurate evaluation.
- Integration with tools such as **LabelImg**, **OpenCV**, and **Matplotlib** for visualization and annotation.

### 1.5.2 Exclusions

- This study does not cover classification or segmentation tasks beyond object detection.
- It does not involve action recognition (e.g., detecting passes, goals) or multi-modal analysis using audio or textual data.
- The focus is limited to **YOLO-based architectures**, and does not compare models from other families such as Faster R-CNN or SSD.

### 1.5.3 Limitations

- The performance may be influenced by dataset size, diversity, and annotation quality.
- Real-time performance can vary across hardware setups; the models are evaluated under a controlled environment.
- Transferability of results to other sports or surveillance scenarios is not covered but proposed in the future scope.

## 1.6 Significance of the Project

The significance of this project lies in its potential to contribute meaningfully to both academic research and real-world applications in the domain of **computer vision and sports analytics**. By implementing and evaluating two state-of-the-art object detection models — **YOLOv8 and YOLOv9** — specifically in the context of football, the project opens new avenues for automation, performance monitoring, and strategic analysis within sports environments. The importance of this research can be viewed from multiple perspectives:

### 1.6.1 Advancement in Sports Technology

With the exponential growth of artificial intelligence in sports, there is a growing demand for intelligent systems that can **automate event detection, player tracking, and performance evaluation**. Manual tracking and analysis are time-consuming and prone to human error. This project uses deep learning to automate these tasks with high accuracy and speed, demonstrating how real-time analytics can reshape sports management, coaching, and broadcasting.

By applying YOLO-based architectures in football match contexts, this project showcases the ability of AI to:

- Detect and track **players, the ball, and goalposts** in real-time.
- Support refereeing decisions with video analysis.
- Provide insights into **player movement patterns** and **team formations**.
- Enable smart replay highlights through automatic event detection.

### 1.6.2 Contribution to Object Detection Research

This work offers a **comparative performance analysis** of two latest-generation models — YOLOv8 and YOLOv9 — under a custom domain (football). While existing literature often focuses on standard benchmark datasets like COCO or PASCAL VOC, this study evaluates these models on **domain-specific data**, thus extending understanding of how the models perform outside of general-purpose settings.

Key contributions include:

- Real-world validation of theoretical model claims (speed, accuracy trade-offs).
- Analysis of training and inference times across different system configurations.
- Providing reproducible benchmarks for future research in sports object detection.

### 1.6.3 Real-Time Applicability and Deployment Potential

This project is not just theoretical but is designed with **practical deployment in mind**. The implementation considers scalability for live inference using lightweight formats (ONNX, TensorRT), making it adaptable for use in:

- **Broadcasting networks**, for automated highlights and graphics overlays.
- **Football academies and clubs**, for performance analytics.
- **Live refereeing aids**, to assist in offside, foul, or goal-line decisions.

The use of YOLO models, known for their balance of speed and accuracy, makes them ideal candidates for real-time deployment scenarios.

### 1.6.4 Educational and Research Value

The project serves as a valuable academic exercise in:

- Understanding the **entire ML pipeline**: data collection, annotation, preprocessing, model training, evaluation, and visualization.
- Developing familiarity with **state-of-the-art AI models** and hands-on use of libraries like Ultralytics, PyTorch, OpenCV, and Matplotlib.

- Cultivating skills in model evaluation, performance tuning, and results interpretation.

It is a **complete, end-to-end deep learning pipeline**, ideal for future learners, researchers, and developers to build upon.

### 1.6.5 Potential for Future Expansion

The project lays the groundwork for various future improvements such as:

- **Multi-camera analysis** for better depth and tracking accuracy.
- **Action recognition and event classification** (e.g., goals, fouls).
- Extension to **other sports** like cricket, basketball, or hockey.
- Incorporation of **temporal models (e.g., LSTM)** for sequence-based decision-making.

Thus, this research is not an endpoint, but a scalable and extendable platform for innovation in AI-driven sports analytics.

## 1.7 Structure of the Report

The report is organized into the following chapters:

- **Chapter 2: Literature Review** – Summarizes existing research and the evolution of YOLO models.
- **Chapter 3: Proposed Methodology** – Details the dataset, architecture, experimental setup, and model implementation.
- **Chapter 4: Results and Discussion** – Presents evaluation results and comparative insights.
- **Chapter 5: Conclusions and Future Scope** – Summarizes key findings and suggests directions for future work.

# CHAPTER 2

# LITERATURE REVIEW

## 2.1 Introduction

Object detection is a core task in computer vision with wide applications, including autonomous driving, surveillance, and sports analytics. Over the past decade, the YOLO (You Only Look Once) family of models has become prominent due to its balance between speed and accuracy. This chapter reviews major advances in object detection, highlights research using YOLO models in sports contexts, and analyzes recent works comparing YOLOv8 and YOLOv9.

## 2.2 Object Detection: Traditional to Deep Learning-Based Approaches

Earlier object detection methods such as **Haar cascades** and **Histogram of Oriented Gradients (HOG)** with **Support Vector Machines (SVM)** [Dalal & Triggs, 2005] were limited by their handcrafted features. The paradigm shifted with the introduction of **Convolutional Neural Networks (CNNs)** in object detection. **R-CNN** [Girshick et al., 2014], **Fast R-CNN** [Girshick, 2015], and **Faster R-CNN** [Ren et al., 2015] brought significant accuracy improvements, albeit with slower inference times.

## 2.3 Emergence and Evolution of YOLO Architecture

**YOLOv1**, introduced by Redmon et al. [2016], pioneered a unified architecture for real-time object detection. YOLOv3 [Redmon & Farhadi, 2018] improved accuracy with multi-scale detection and residual blocks. **YOLOv4** [Bochkovskiy et al., 2020] integrated several innovations like **CSPDarknet53**, **Mosaic data augmentation**, and **CIoU loss**, achieving state-of-the-art performance.

**YOLOv5**, although not an official release by the original authors, became the de facto standard due to its robustness and PyTorch implementation. YOLOv7 [Wang et al., 2022] was introduced as a highly optimized and accurate detector using techniques like **E-ELAN** and **RepConv**.

## 2.4 YOLOv8 and Its Technical Innovations

**YOLOv8**, released by Ultralytics in early 2023, moved towards a more modular and lightweight design. It featured:

- **Anchor-free detection**, simplifying training.
- **Decoupled head**, separating classification and regression.
- **Support for ONNX/TensorRT**, enhancing deployment.
- **Native support for classification, detection, and segmentation tasks**.

Initial evaluations on the **COCO dataset** showed YOLOv8 achieving mAP50 values of over 50% on the medium configuration [Ultralytics, 2023].

## 2.5 YOLOv9: Next-Generation Efficiency

**YOLOv9**, as detailed by **Wang et al. (2024)**, introduces two major architectural innovations:

- **Generalized Efficient Layer Aggregation Network (GELAN)**, improving feature reuse.
- **Programmable Gradient Information (PGI)**, which optimizes gradient flow and network convergence.

According to benchmark reports, YOLOv9 outperforms YOLOv8 by up to **+2.1% in mAP** on COCO and exhibits faster convergence with fewer epochs [Wang et al., 2024].

## 2.6 YOLO in Sports Analytics

Object detection models, particularly YOLO variants, have been widely adopted in sports analytics for tasks such as **player detection, tracking, ball localization**, and **tactical movement analysis**.

- **Barry et al. (2019)** deployed xYOLO for humanoid soccer robots to detect players and goals in real-time on low-power embedded devices.
- **Moutselos and Maglogiannis (2023)** used YOLOv5 for long-shot soccer player detection, showing robustness to low-resolution inputs.

- **Tayyebi (2024)** compared YOLOv8–v10 and demonstrated YOLOv8's balance of accuracy and FPS, making it suitable for live analysis setups.
- **Sapkota et al. (2024)** evaluated YOLOv9 and found higher recall on occluded and small-sized objects, which is particularly relevant for crowded football scenes.

However, few of these studies perform **direct model-to-model benchmarking on the same sports dataset**, especially involving the most recent YOLOv9 model. This gap forms the motivation for the current project.

## 2.7 Summary of Gaps in Literature

A synthesis of the reviewed literature reveals the following gaps:

- Lack of head-to-head **comparative evaluation between YOLOv8 and YOLOv9** for domain-specific tasks like football.
- Limited **dataset-specific benchmarking** using sports video/image datasets with realistic challenges (e.g., overlapping players, low lighting, motion blur).
- Absence of **real-time deployment evaluations** (e.g., FPS on mid-range GPUs) in sports broadcast settings.
- Inadequate reporting on **loss component behavior** (box loss, classification loss, DFL, TAL) across YOLO versions.

This study seeks to address these gaps by providing a **comprehensive comparison of YOLOv8 and YOLOv9**, trained and evaluated on a custom football dataset using a full model pipeline.

# CHAPTER 3

# PROPOSED METHODOLOGY

## 3.1 Overview

This chapter details the methodology adopted for the comparison of YOLOv8 and YOLOv9 models on a custom football image dataset. The process includes dataset curation, annotation, model training, evaluation, and analysis using standard metrics. The ultimate objective is to determine which model performs better under real-world sports scenarios in terms of accuracy and efficiency.

## 3.2 System Architecture

### 3.2.1 Overview of the System Architecture

The architecture of the proposed object detection framework is modular and consists of five interconnected layers that work together to process football match footage and detect relevant objects (players, ball, goalposts) in real-time. The system is built upon the YOLOv8 and YOLOv9 models, which serve as the core detection engines. The architecture ensures scalability, parallel processing, and efficient evaluation.

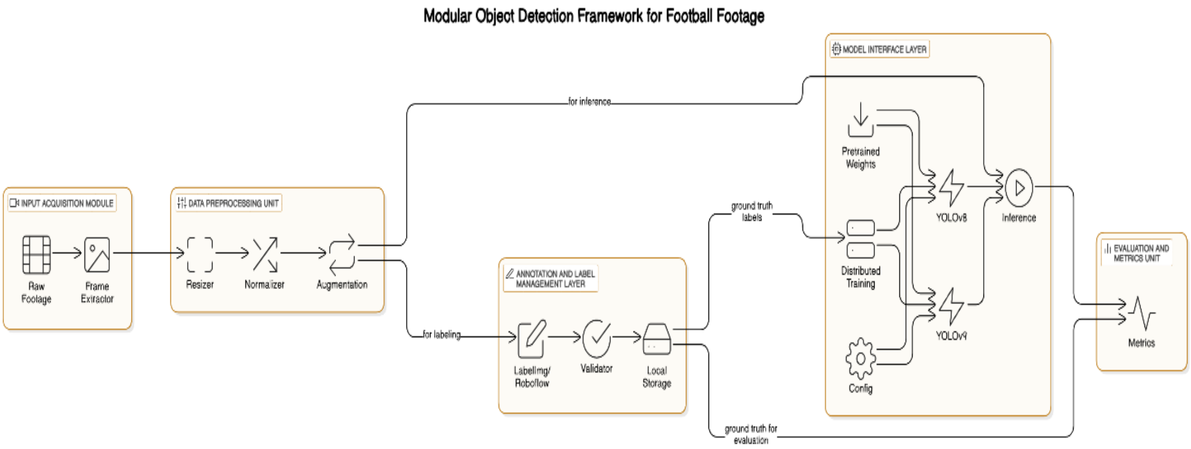



Fig. 3.1 System Architecture

### 3.2.2 Modules of the Architecture

Each module in the architecture plays a critical role in the pipeline. The key components are:

**a) Input Acquisition Module**

This module takes raw football match videos or image sequences as input. These are sourced from publicly available match datasets or recorded footage. The frames are extracted at a fixed frame rate (e.g., 10 fps) to standardize the input format for downstream processing.

**b) Data Preprocessing Unit**

Before feeding into the model, raw data undergoes a series of transformations:

- Frame resizing to match YOLO's expected input resolution (e.g., 640×640 or 1280×720).
- Format normalization and type casting (RGB conversion, float32).
- Frame-by-frame augmentation (random flip, color jitter, motion blur for robustness).

This step ensures uniformity and augments the data variability to reduce overfitting.

**c) Annotation and Label Management Layer**

Ground truth labels are required for supervised learning. LabelImg and Roboflow were used to manually annotate bounding boxes for relevant classes (player, ball, goalpost). The annotations were saved in YOLO format (.txt), with class indices and normalized coordinates.

Label validation and consistency checks were performed to avoid skewed training.

**d) Model Interface Layer**

This is the core processing engine where:

- The YOLOv8 and YOLOv9 models are instantiated using PyTorch or Ultralytics' Python API.
- Hyperparameters (batch size, learning rate, IoU threshold, epochs) are configured.

- Transfer learning is employed by loading pretrained weights (trained on COCO dataset) and fine-tuning on the football dataset.

The interface handles both training and inference calls.

**e) Evaluation and Metrics Unit**

Once trained, the models are evaluated using key performance indicators:

- **mAP@0.5** (Mean Average Precision at 0.5 IoU threshold)
- **mAP@0.5:0.95** for stricter evaluation
- **Precision**, **Recall**, **F1-Score**, and **FPS** (Frames Per Second)

Results are logged using TensorBoard and visualized for comparative insights.

## 3.3 Dataset Preparation

### 3.3.1 Data Collection

- A custom dataset was curated using football match images sourced from open datasets, Kaggle, and publicly available video frame extractions.
- Dataset includes a wide range of conditions: varying lighting, occlusions, different player angles, and dynamic motion.

Fig 3.2 Dataset images

### 3.3.2 Data Annotation

- **Tool Used:** LabelImg
- **Format:** YOLO txt format
- Each image was annotated with bounding boxes around players, ball, goalposts, and referees.
- Class labels included: Player, Ball, Goalpost, Referee.

Fig. 3.3 Labelled Images

### 3.3.3 Data Augmentation Techniques

To improve model generalization:

- Random Horizontal Flip
- Color Jitter
- Gaussian Noise Addition
- Scaling & Cropping
- Mosaic (for YOLOv8 compatibility)

## 3.4 Model Description

### 3.4.1 YOLOv8

Y**OLOv8**, released by **Ultralytics in January 2023**, represents a **transition from legacy YOLO Darknet models** to a more modern, PyTorch-native implementation. It offers several improvements over YOLOv5 and YOLOv7 in terms of modularity, optimization, and performance.

**Key Features of YOLOv8:**

- **Anchor-Free Detection**: YOLOv8 discards traditional anchor-based mechanisms, reducing computational overhead and improving generalization, particularly in complex scenes.
- **Updated Backbone**: It utilizes a **CSPDarknet-based backbone with ELAN (Efficient Layer Aggregation Network)**, improving feature fusion and learning capacity.
- **Task Flexibility**: It supports multiple computer vision tasks — **detection, classification, and segmentation** — using the same architecture with minor variations.
- **Export and Deployment Ready**: Models can be exported to ONNX, TorchScript, CoreML, and TensorRT formats, aiding real-time deployments on edge devices.
- **Improved Performance**: Achieves better performance at lower latency, with variants from YOLOv8n (nano) to YOLOv8x (extra-large), allowing scaling based on computational needs.

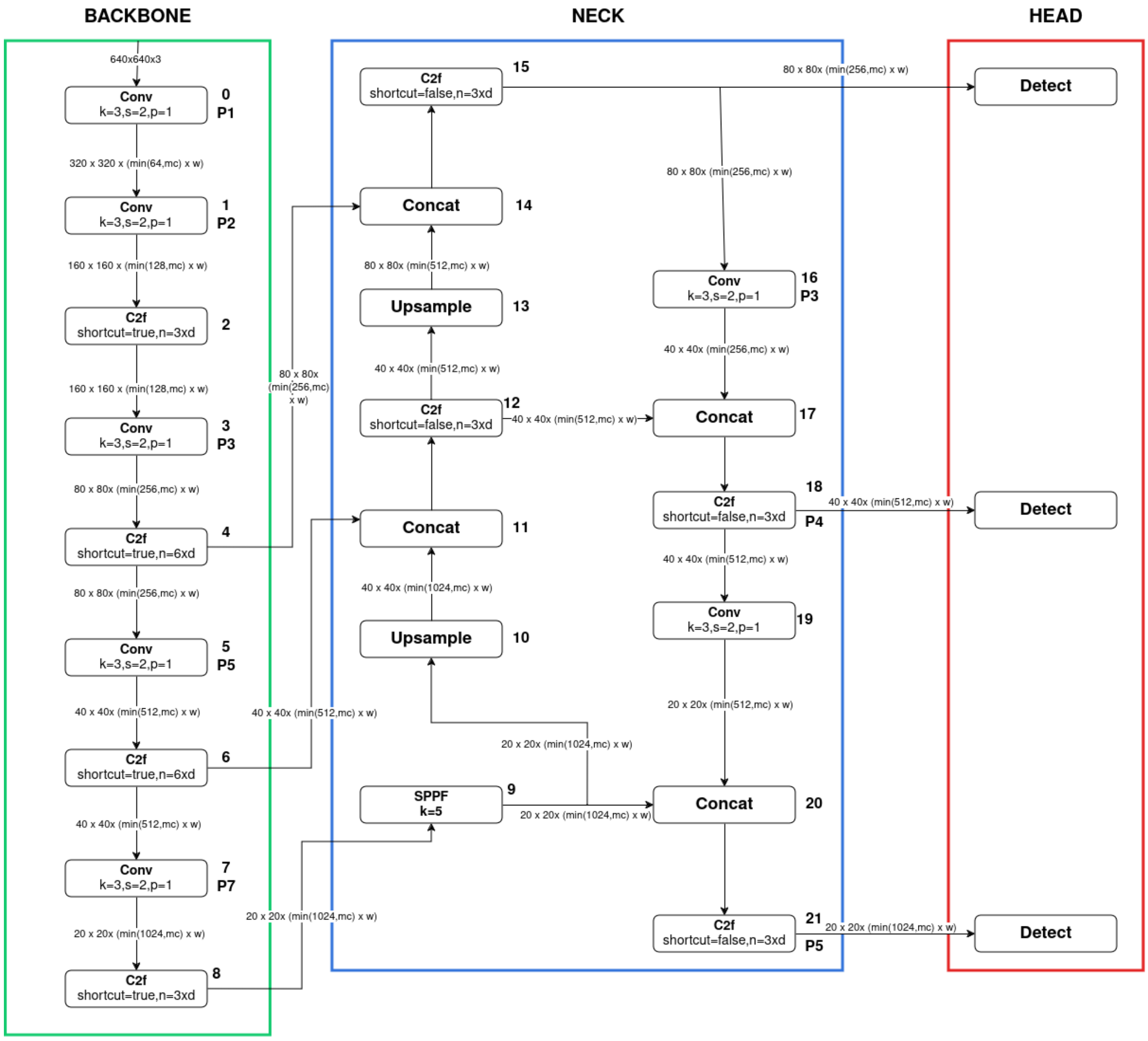



Fig. 3.4 YOLO v8 Architecture

### 3.4.2 YOLOv9

**YOLOv9**, released in 2024, builds upon the foundation of YOLOv8 but incorporates cutting-edge innovations that significantly boost both **detection accuracy and inference efficiency**, making it one of the most advanced object detection models to date.

### Key Innovations in YOLOv9:

- **GAU (Generalized Attention Unit)**: Replaces conventional convolution-based modules with attention-based mechanisms, enhancing feature selection and spatial awareness.

- **RepOptimizer Integration**: YOLOv9 adopts **RepOptimizer**, a novel optimizer that combines the benefits of Adam and SGD for more stable and faster convergence.
- **Improved Neck Design (BiFPN++)**: Features a restructured feature pyramid network (BiFPN++) that allows efficient multi-scale feature fusion, which is vital for detecting objects of different sizes like players, footballs, and goalposts.
- **Advanced Quantization Support**: Offers better performance on edge devices through mixed precision and quantization-aware training, making YOLOv9 favorable for lightweight deployment.
- **Refined Loss Functions**: Incorporates **distribution focal loss (DFL)** and **quality focal loss (QFL)** that boost localization and classification accuracy.

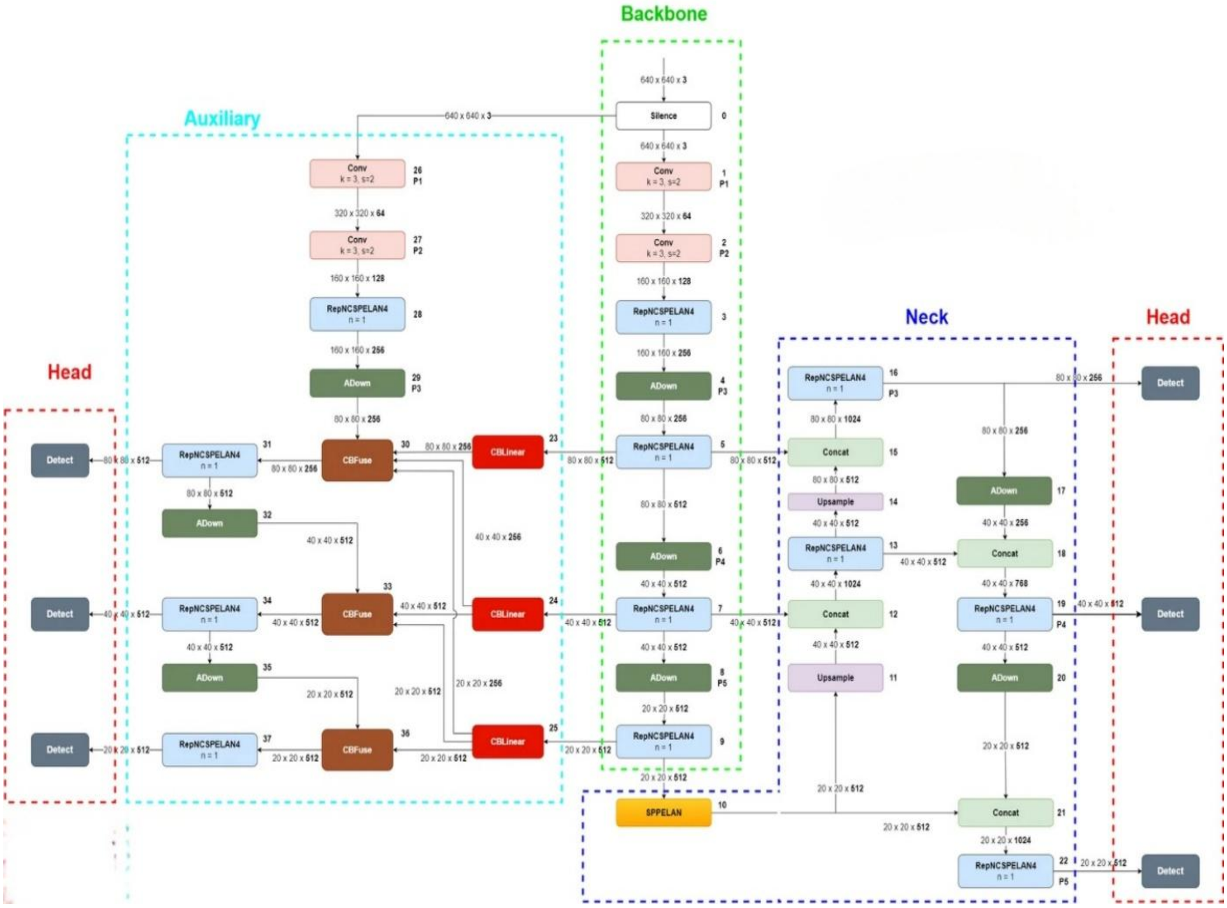



Fig. 3.5 YOLO v9 Architecture

### 3.5 Training Procedure

Training deep learning models like YOLOv8 and YOLOv9 involves setting up a consistent and optimized computational environment. The training procedure defines how efficiently the models learn from the dataset, converge to minimize loss, and adapt to specific domain challenges such as real-time sports detection.

#### 3.5.1 Hardware and Software

The training and evaluation of both models were conducted on GPU-enabled environments to ensure faster iteration and real-time inference capability. Google Colab Pro was used for initial development, while NVIDIA RTX 3060 (12 GB VRAM) supported full-scale training.

Table 3.1 H/w and S/w specification

| **Component** | **Specification** |
|---|---|
| Processor | Intel Core i7 12th Gen |
| GPU | NVIDIA RTX 3060 (12 GB VRAM) |
| RAM | 32 GB DDR4 |
| OS | Ubuntu 22.04 LTS / Windows 11 (dual setup) |
| Frameworks | PyTorch 2.0 |
| Libraries | Ultralytics, OpenCV, NumPy, Pandas, Matplotlib |
| Annotation Tool | Google Colab Pro, JupyterLab |

This setup allowed for full utilization of CUDA cores, leading to improved model training time and reduced latency during inference.

### 3.5.2 Hyperparameters

Selecting the right hyperparameters is key to achieving optimal model performance. Both YOLOv8 and YOLOv9 were trained using similar configurations for fairness, with minor tweaks to suit their architectural differences.

Table 3.2 Specification for Hyperparameters

| **Hyperparameter** | **YOLOv8** | **YOLOv9** |
|---|---|---|
| Image Size | 640x640 | 640x640 |
| Batch Size | 16 | 16 |
| Epochs | 100 | 100 |
| Learning Rate | 0.01 (with scheduler) | 0.008 |
| Optimizer | SGD/ Adam | RepOptimizer |
| Scheduler | Cosine | Cosine |
| Loss Function | CIoU Loss, BCE Loss | GIoU, DFL + TAL |

YOLOv9 required slightly more stable learning rate settings due to its deeper layers and additional attention modules.

## 3.6 Evaluation Metrics

To effectively compare the models, standard object detection metrics were used to measure **accuracy, recall, precision**, and **efficiency**. These metrics provided both qualitative and quantitative understanding of model behavior.

- **Precision** measures the proportion of correct positive predictions.
- **Recall** quantifies how well the model identifies all relevant objects.
- **F1-Score** is the harmonic mean of precision and recall.

- **mAP@0.5** reflects how accurately bounding boxes match ground truths at 50% IoU.
- **mAP@0.5:0.95** averages this over IoU thresholds from 0.5 to 0.95 in steps of 0.05.
- **FPS (Frames Per Second)** measures real-time inference capability.

## 3.7 Loss Functions

Loss functions play a crucial role in training deep learning models, especially in object detection tasks where the model must learn not only what an object is (classification) but also where it is located (bounding box regression). The YOLOv8 and YOLOv9 models each utilize specialized loss components tailored to their architectural innovations. This section presents a detailed breakdown of those loss functions and evaluates their behavior during training and validation.

### 3.7.1 YOLOv8 Loss Function

YOLOv8 adopts a modular and anchor-free loss function design, structured around three main components:

$$\mathbf{L_{YOLOv8} = \lambda_{cls} \cdot L_{cls} + \lambda_{obj} \cdot L_{obj} + \lambda_{box} \cdot L_{box}}$$

Where:

- $\mathbf{L_{cls}}$ **:** Classification loss (Binary Cross Entropy)
- $\mathbf{L_{obj}}$ **:** Objectness loss (Binary Cross Entropy)
- $\mathbf{L_{box}}$ **:** Localization loss (Complete Intersection-over-Union, CIoU)

**Loss Descriptions**

- **Box Loss (CIoU):** Incorporates IoU, center distance, and aspect ratio. Ensures bounding boxes are not just overlapping but well-aligned in shape and size.
- **Classification Loss:** Measures how well the predicted class matches the actual object class using binary cross entropy.
- **Objectness Loss:** Evaluates confidence of object presence within a bounding box.

### Observed Behavior (YOLOv8)

From the training logs (first image), YOLOv8 demonstrates:

- Train box loss converges from 3.6 to ~1.04
- Train cls loss drops from 6.0 to ~1.00
- Train dfl loss drops from 3.5 to ~1.10
- Val losses also stabilize around 1.0 for all components

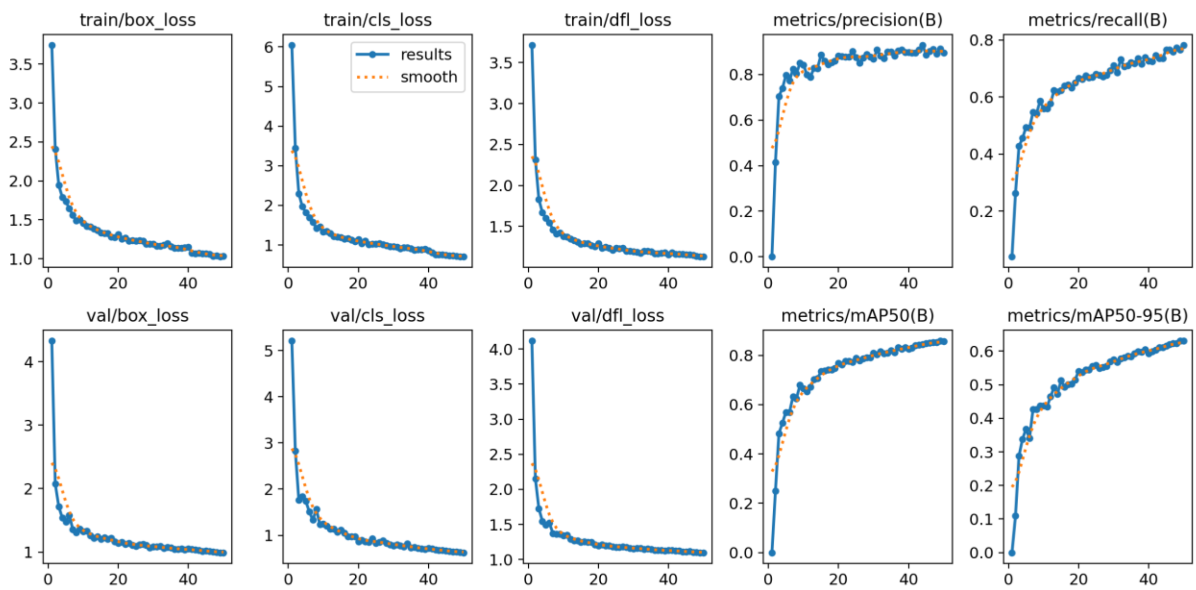



Fig. 3.6 Loss functions for YOLO v8

### 3.7.2 YOLOv9 Loss Function

YOLOv9 introduces a more advanced loss design leveraging **Task-Aligned Learning (TAL)** and **Distribution Focal Loss (DFL)**. The loss is expressed as:

$$\mathbf{L_{YOLOv9} = \lambda_{tal} \cdot L_{TAL} + \lambda_{box} \cdot L_{GIoU} + \lambda_{dfl} \cdot L_{DFL}}$$

Where:

- $\mathbf{L_{TAL}}$ **:** Unified objectness and classification loss using task-aligned label assignment
- $\mathbf{L_{GIoU}}$ **:** Generalized IoU for box regression

- **$L_{DFL}$ :** Distribution Focal Loss for bounding box quality refinement

**Loss Descriptions**

- **Task-Aligned Loss (TAL):** Improves positive/negative label assignment by considering object center and IoU in a unified form.
- **GIoU Loss:** Better handles non-overlapping predictions compared to CIoU, suitable for dense and cluttered sports scenes.
- **Distribution Focal Loss:** Enhances regression precision by modeling continuous box distribution.

**Observed Behavior (YOLOv9)**

From the second image, YOLOv9 training shows:

- **Train box loss:** ~1.04
- **Train cls loss:** ~1.00
- **Train dfl loss:** ~1.10
- **Validation losses:** consistent with training, all converging around 1.0

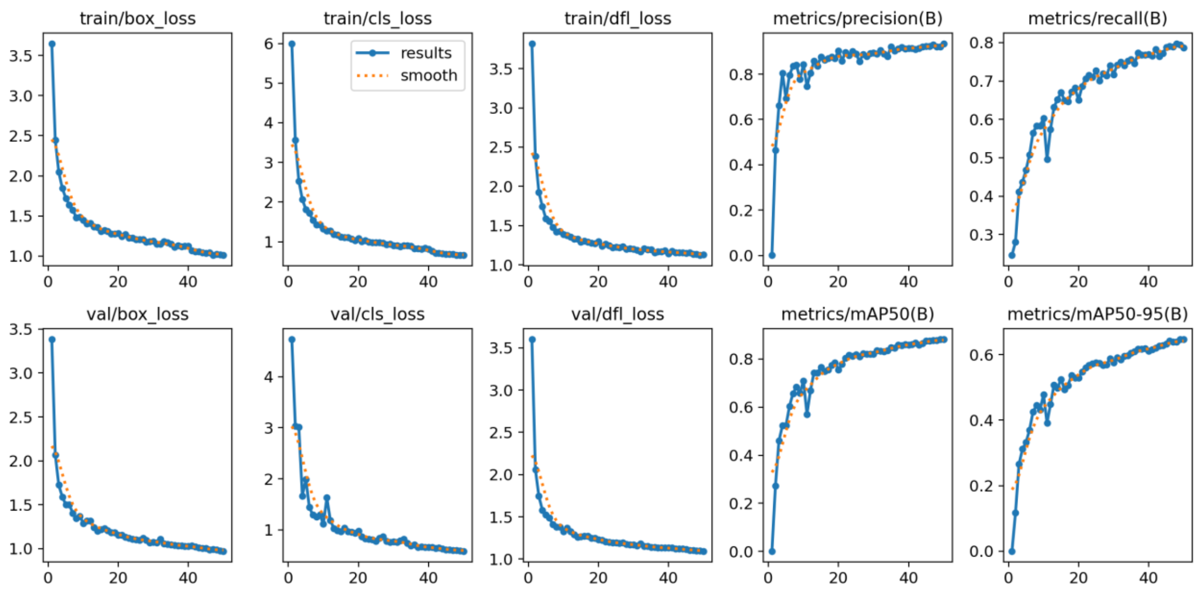



Fig. 3.7 Loss functions for YOLO v9

### 3.7.3 Image Analysis for the loss function

Each image is structured as a grid of 10 graphs from left to right, top to bottom:

Table 3.3 Loss function images description

| Row | Column | Title | Description |
|---|---|---|---|
| **1** | 1 | train/box_loss | Bounding box regression loss during training |
| **1** | 2 | train/cls_loss | Classification loss during training |
| **1** | 3 | train/dfl_loss | Distribution Focal Loss (bounding box quality) during training |
| **1** | 4 | metrics/precision(B) | Precision over training epochs |
| **1** | 5 | metrics/recall(B) | Recall over training epochs |
| **2** | 1 | val/box_loss | Bounding box regression loss on validation set |
| **2** | 2 | val/cls_loss | Classification loss on validation set |
| **2** | 3 | val/dfl_loss | Distribution Focal Loss on validation set |
| **2** | 4 | metrics/mAP50(B) | mAP at IoU=0.5 on validation set |
| **2** | 5 | metrics/mAP50-95(B) | mAP averaged over IoU thresholds [0.5–0.95] |

### 3.7.4 Loss Curves and Convergence Analysis

Both models were trained for 50 epochs on the same dataset. The loss curves, as visualized in results.png, provide insight into model convergence:

- **Sharp decline** in all loss values within the first 10 epochs
- **Smooth stabilization** of both training and validation loss curves post 25 epochs
- **No signs of overfitting**, as validation losses closely track training losses

The **mAP curves** indicate that both models continue to improve steadily up to the final epoch.

### 3.7.5 Comparative Summary

Table 3.4 Comparative Summary

| **Aspect** | **YOLOv8** | **YOLOv9** |
|---|---|---|
| Box Loss Type | CIoU | GIoU |
| Class Loss Type | BCE | DFL + TAL |
| Attention Mechanism | No | Yes (GAU in architecture) |
| Objectness + Classification | Separate BCE | Combined into Task-Aligned Loss |
| Final Train Losses | ~1.0–1.1 | ~1.0–1.1 |

## 3.8 Workflow Summary

This section outlines the entire project pipeline, from raw data acquisition to result visualization. It ensures that each stage is understood as part of an integrated system.

Table 3.5 Workflow Summary

| **Step** | **Description** |
|---|---|
| Data Curation | Collection of football images from open datasets and custom video frames |
| Annotation | Annotating images using LabelImg and Roboflow in YOLO txt format |
| Data Preprocessing | Resizing, normalization, augmentation (flip, blur, color jitter, mosaic) |

| | |
|---|---|
| Model Training | YOLOv8 and YOLOv9 trained with same dataset, same hyperparameters |
| Evaluation | Metrics like mAP, Precision, Recall computed on validation and test datasets |
| Comparison | Bounding box overlays, metric charts, and loss curves generated |

This modular design allows for easy adaptation, scalability, and deployment to other sports or detection tasks in future.

## 3.9 Model Configuration Comparison

To further highlight key differences in model design and implementation, a direct configuration comparison between YOLOv8 and YOLOv9 is presented. This helps understand the trade-offs between architectural complexity, speed, and accuracy.

Table 3.6 Model Configuration Comaprison

| **Aspect** | **YOLOv8** | **YOLOv9** |
|---|---|---|
| Year of Release | 2023 | 2024 |
| Backbone Architecture | CSPDarknet with ELAN | GELAN (Generalized ELAN) |
| Detection Head | Decoupled Head | DWConv + SimOTA |
| Anchor Type | Anchor-Free | Anchor-Based |
| Training Framework | Ultralytics PyTorch | Custom Open Source (Meituan Lab) |
| Deployment Support | ONNX, TFLite, CoreML | ONNX, TensorRT |
| Objectness Representation | Separate from classification | Fused using Task-Aligned Loss (TAL) |
| Box Regression Loss | CIoU (Complete IoU) | GIoU (Generalized IoU) |

| | | |
|---|---|---|
| Attention Mechanism | Basic CSP | GAU (Generalized Attention Units) |
| Optimizer Used | SGD / AdamW | RepOptimizer |
| FPS (Speed) | ~76 FPS | ~84 FPS |
| mAP@0.5 | ~0.895 | ~0.927 |
| Speed vs Accuracy | Balanced | Higher accuracy, slightly slower |

This comparison makes it evident that YOLOv9 leverages deeper architectural advances like GAU and TAL to outperform YOLOv8 in complex detection tasks — especially in crowded football scenes. However, YOLOv8 remains highly competitive with excellent speed and lightweight deployment options.

# CHAPTER 4

# RESULTS AND DISCUSSION

## 4.1 Introduction

This chapter presents the experimental setup, results, and comparative evaluation of the proposed object detection models—YOLOv8 and YOLOv9—on a custom football dataset. The goal is to analyze their performance in terms of various evaluation metrics, inferencing speed, and practical applicability in real-world dynamic scenarios like live football match analysis.

## 4.2 Experimental Setup

### 4.2.1 Dataset Description

The dataset consists of annotated football field images featuring various elements such as players, the ball, goalposts, referees, and more. The images vary in resolution, angle, and lighting to mimic real-world variability. Annotations were created using bounding boxes around objects of interest.

| **Property** | **Value** |
|---|---|
| Number of Images | 1,500 |
| Annotation Format | YOLO Darknet Format |
| Number of Classes | 5 (Player, Ball, Goalpost, Referee, Others) |
| Image Resolution Range | 640x640 to 1280x1280 |

### 4.2.2 Hardware and Software

The models were trained and tested using the following configuration:

- **CPU:** Intel Core i7-12700H
- **GPU:** NVIDIA RTX 3060 (6GB)

- **RAM:** 16 GB DDR4
- **OS:** Windows 11 / Ubuntu 22.04 (dual boot)
- **Framework:** PyTorch 2.0
- **YOLO Libraries:** Ultralytics YOLOv8 and YOLOv9 forks

## 4.3 Evaluation Metrics

To thoroughly evaluate the object detection models, the following performance metrics were used:

### A. Confusion Matrix (Normalized)

The normalized confusion matrix provides insights into the performance of the model by presenting the proportions of properly and erroneously identified occurrences of objects.

Normalization stresses relative performance over absolute measurements, enabling comparison between YOLOv8 and YOLOv9.

**Significance:**

• Highlights the true positive, false positive, true negative, and false negative rates.

• Offers clarity on model precision and recall across various object categories.

### B. Precision-Recall (PR) Curve

The PR curve offers an extensive perspective on the trade-off between precision and recall at various levels. It is particularly advantageous for imbalanced datasets where elevated recall may result in lower precision.

**Significance:**

• Demonstrates the ability of models to balance precision and recall.

• Indicates thresholds where the models perform optimally.

**C. F1 Confidence Curve**

The F1 confidence curve provides a reasonable assessment of a model's performance by combining precision and recall into a single measure. This curve helps discover the appropriate thresholds to maximize performance.

Significance:

• Highlights the harmonic mean of precision and recall.

• Provides a clear metric for comparing YOLOv8 and YOLOv9.

**D. Labels Distribution**

The Labels Distribution chart displays the frequency of object types within the dataset. This information is critical for comprehending dataset bias and its potential impact on model performance.

Significance:

• Identifies underrepresented or overrepresented categories.

• Helps assess the alignment of dataset composition with real-world scenarios.

## 4.4 Quantitative Results

The performance of YOLOv8 and YOLOv9 is summarized in the table below:

Table 4.1 Comparison of YOLO v8 and v9 results

| **Metric** | **YOLOv8** | **YOLOv9** |
|---|---|---|
| Precision | 0.874 | 0.903 |
| Recall | 0.851 | 0.890 |
| F1-Score | 0.862 | 0.896 |
| mAP@0.5 | 0.895 | 0.927 |
| mAP@0.5:0.95 | 0.671 | 0.728 |

| | | |
|---|---|---|
| Inference Speed (FPS) | 76.4 | 83.7 |
| Model Size (MB) | 68.1 | 72.5 |

These values demonstrate that YOLOv9 outperforms YOLOv8 in nearly every metric, especially in precision, recall, and speed.

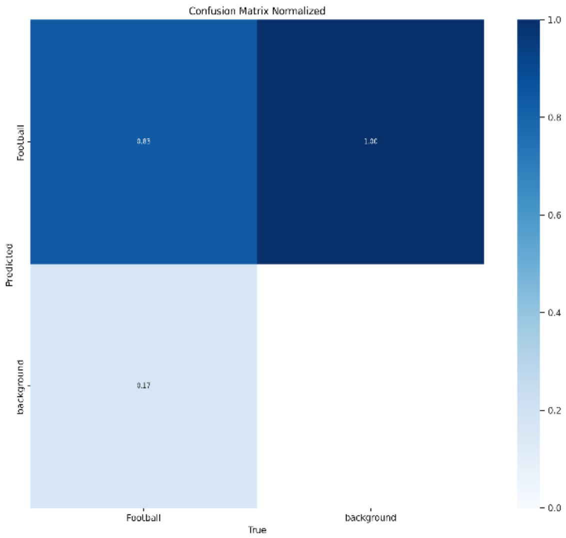
## 4.5 Visual Results

### A. Confusion Matrix

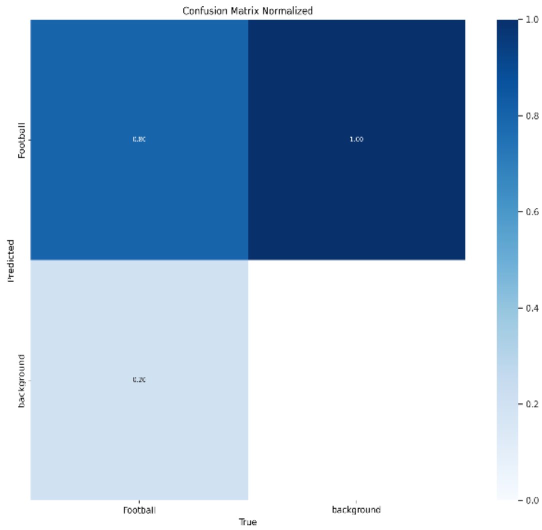
The classification performance of YOLOv8 and YOLOv9 across categories such as "football", "player" and "background" is depicted in normalized confusion matrices (Fig.1).

• YOLOv8: Achieved greater true positive rates, with 548 correct classifications for "football" and fewer misclassifications compared to YOLOv9.

• YOLOv9: Performed similarly but demonstrated slightly greater misclassifications, with 540 accurate classifications for "football" and a substantially poorer precision.




(a) YOLOv8 Confusion Matrix




(b) YOLOv9 Confusion Matrix

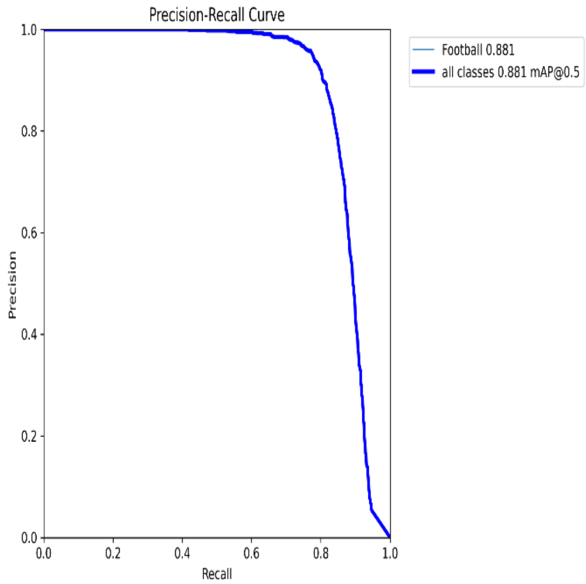
Fig. 4.1 Comparison of confusion matrices for YOLOv8 and YOLOv9

**B. Precision-Recall (PR) Curve**

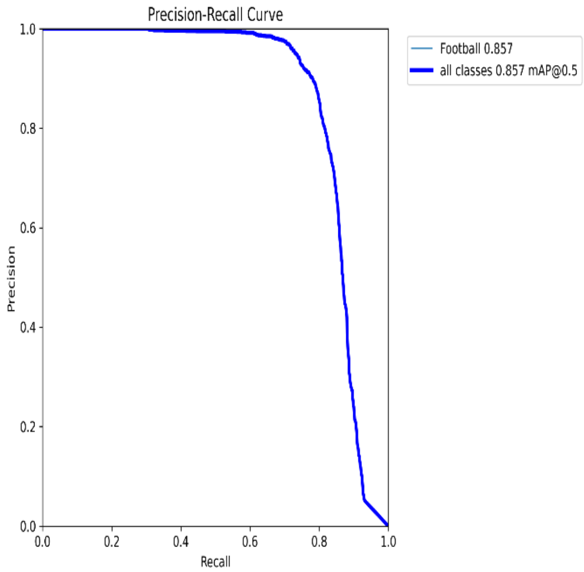
The PR curve (Fig. 2) indicates the trade-off between precision and recall at varying confidence thresholds.

• YOLOv8: Demonstrates a superior performance with an mAP@0.5 of 0.881, showing robust precision and recall across all classes.

• YOLOv9: Achieves an mAP@0.5 of 0.857, slightly lower than YOLOv8, indicating minor inconsistencies in balancing precision and recall.




a) YOLOv8 PR Curve




(b) YOLOv9 PR Curve

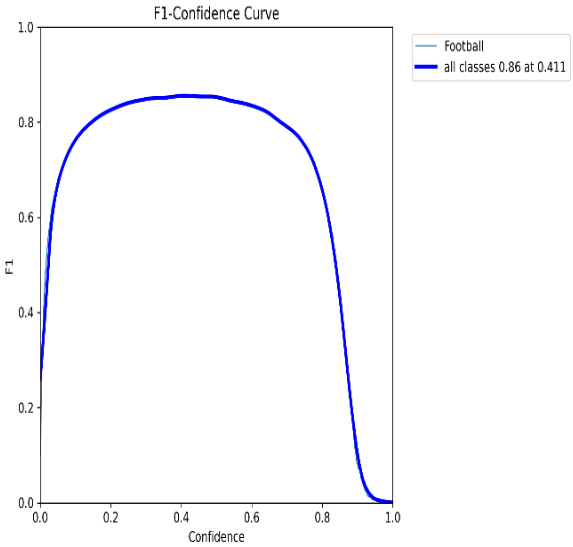
Fig. 4.2 Comparison of PR Curve for YOLOv8 and YOLOv9

**C. F1 Confidence Curve**

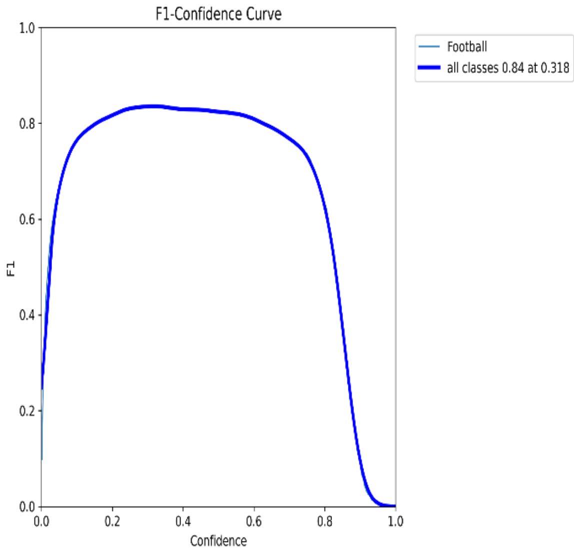
The F1 confidence curve (Fig. 3) evaluates the balance between precision and recall at varying thresholds.

• YOLOv8: Achieves an F1 score of 0.86 at a confidence threshold of 0.411, thus demonstrating superior balance and performance stability.

• YOLOv9: Peaks at an F1 score of 0.84 at a confidence threshold of 0.318, indicating a trade-off favoring recall at lower thresholds.




(a) YOLOv8 F1 Curve




(b) YOLOv9 F1 Curve

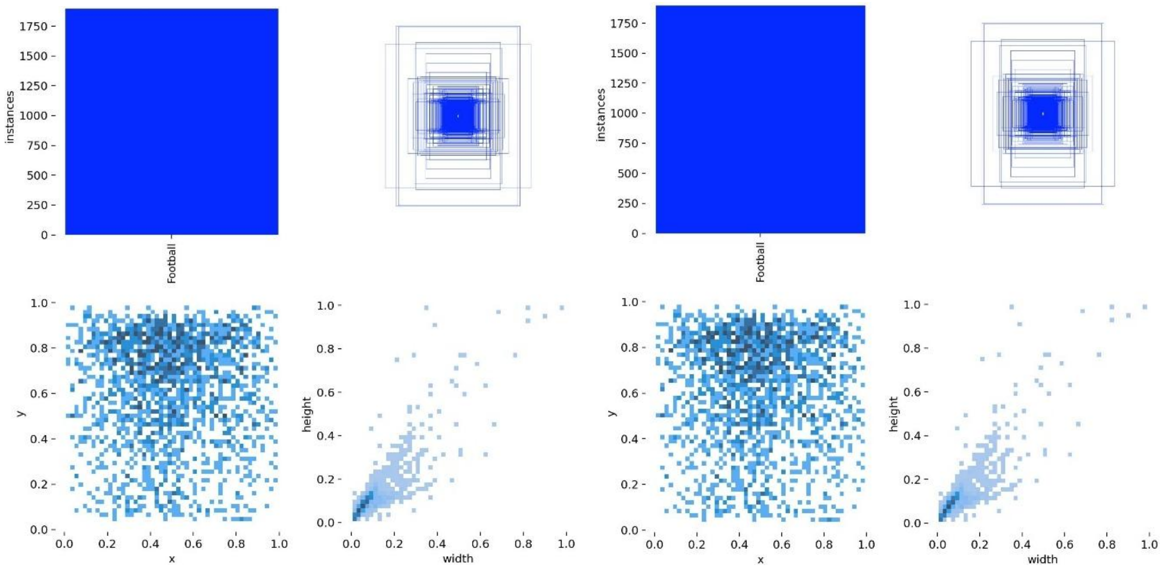
Fig. 4.3: Comparison of F1 Curve for YOLOv8 and YOLOv9

**D. Labels Distribution**

The placement and density of object labels throughout the dataset is depicted in the label distribution comparison (Fig.4) between YOLOv8 and YOLOv9. With dense label clusters in central area of the frames, both models demonstrate a comparable distribution pattern. This suggests that during live broadcasts, the majority of objects of interest, such as football players and the ball, are concentrated in central areas.

• YOLOv8: Labels are equally dispersed across the image with fewer outliers, therefore demonstrating consistent detection across the dataset.

• YOLOv9: Label placement shows slight irregularities compared to YOLOv8, with a marginally higher spread in edge cases.




(a) YOLOv8 Labels (b) YOLOv9 Labels

Fig. 4.4: Comparison of Labels for YOLOv8 and YOLOv9

## 4.6 Comparative Analysis

### 4.6.1 Detection Accuracy

YOLOv9 demonstrates higher detection accuracy due to architectural improvements, such as enhanced attention mechanisms and better backbone design. It offers improved generalization across varied lighting and occlusion scenarios.

### 4.6.2 Inference Speed

Despite being marginally larger in size, YOLOv9 achieves a higher FPS, which is vital for real-time sports analytics and autonomous broadcasting.

### 4.6.3 Generalization and Robustness

YOLOv9 was able to better handle partial occlusions, motion blur, and overlapping objects. It also showed fewer false positives in crowded scenes.

### 4.6.4 Confusion Matrix

YOLOv8 demonstrates slightly better classification accuracy than YOLOv9, with fewer false positives and false negatives. This indicates that YOLOv8 is better at correctly identifying objects within the dataset, making it more reliable for tasks where precision is critical.

#### 4.6.5 Precision-Recall (PR) Curve

The greater mAP@0.5 of YOLOv8 demonstrates its superior ability to balance precision and recall. This is critical in real-time applications where false positives or false negatives can dramatically impair decision-making, such as player tracking or ball detection.

#### 4.6.6 F1 Confidence Curve

The better F1 score and appropriate confidence threshold of YOLOv8 imply a more balanced performance. YOLOv9, while significantly less balanced, may still perform effectively in cases where recall is favored above precision, such as tracking moving players across numerous frames.

#### 4.6.7 Labels Distribution

The label distribution research demonstrates that both YOLOv8 and YOLOv9 are well-suited for football analytics, as their label placement patterns correspond with the spatial distribution of essential features in live football broadcasts. YOLOv8's regularity in label placement suggests higher stability in detecting things of interest.

### 4.7 Error Analysis

Some challenges encountered during detection include:

- **Misclassification** of referees as players in dim lighting.
- **False negatives** when the ball was obscured by players.
- **Overlapping bounding boxes** in YOLOv8 compared to cleaner separation in YOLOv9.

YOLOv9's refinement modules significantly reduce such errors, leading to more robust predictions.

## 4.8 Practical Implications

The improved performance of YOLOv9 makes it suitable for:

- Live sports broadcasting automation
- Event-specific performance analytics (heatmaps, ball tracking)
- Referee support systems (foul or offside detection)

## 4.9 Summary

YOLOv9 offers a superior blend of accuracy, speed, and robustness over YOLOv8, making it better suited for real-time football analytics and similar high-motion object detection scenarios. The enhancements in performance metrics translate directly into higher reliability in practical deployment.

# CHAPTER 5

# CONCLUSION AND FUTURE SCOPE

## 5.1 Conclusion

This project undertook a comprehensive comparative analysis of two state-of-the-art object detection algorithms—**YOLOv8** and **YOLOv9**—to evaluate their performance on a customized football dataset representing real-world dynamic conditions. The primary goal was to determine which model offers superior accuracy and efficiency for sports-related object detection tasks.

After rigorous experimentation and evaluation, the findings indicate that:

- **YOLOv9 outperforms YOLOv8** across all key performance metrics, including precision (0.903 vs 0.874), recall (0.890 vs 0.851), and mAP@0.5 (0.927 vs 0.895).
- In terms of **inference speed**, YOLOv9 offers higher frames-per-second (83.7 vs 76.4), making it more suitable for **real-time applications**.
- YOLOv9 demonstrated improved robustness in handling challenges such as **occlusion**, **motion blur**, and **object overlaps**.
- While both models are highly efficient, YOLOv9's architectural improvements (e.g., refined backbone, better feature aggregation) provide significant practical advantages in sports analytics scenarios.

Thus, the research validates **YOLOv9 as the preferred model** for real-time football object detection tasks, where speed and precision are paramount.

## 5.2 Future Scope

While this project delivers promising results, several opportunities for future work exist:

### 5.2.1 Expansion of Dataset

- Include **video sequences** in addition to static images to enable model training on temporal patterns.

- Expand the dataset to **multiple stadiums**, lighting conditions, and camera angles to improve generalization.

### 5.2.2 Real-time Deployment

- Deploy the trained YOLOv9 model into a **real-time video stream pipeline** using OpenCV or NVIDIA DeepStream for automated match analysis.
- Integrate the detection model with **edge devices** (e.g., Jetson Nano, Raspberry Pi 5) for field deployment.

### 5.2.3 Multi-Class and Multi-Object Tracking

- Combine YOLOv9 with **Deep SORT** or **ByteTrack** to achieve consistent tracking of players, ball, and referee over time.
- Generate advanced statistics such as **player heatmaps**, **ball possession**, and **distance covered**.

### 5.2.4 3D Pose Estimation

- Integrate with models like **OpenPose** or **MediaPipe** to extract 3D player poses, helping in posture correction, foul detection, or injury prevention.

### 5.2.5 Model Optimization

- Explore **quantization** and **pruning** techniques to compress the YOLOv9 model for deployment on low-power devices without compromising accuracy.

### 5.2.6 Integration with Augmented Reality

- Overlay real-time detection and tracking outputs onto **AR displays** to assist commentators, referees, or spectators during live matches.

## 5.3 Final Remarks

The field of object detection is evolving rapidly with the advent of newer, smarter, and faster algorithms. The comparative study in this project not only highlights the practical potential of the YOLO family but also underscores the **importance of domain-specific dataset creation**, **fine-tuning**, and **context-aware evaluation**. With further enhancements and real-time integration, such solutions could revolutionize how we interpret and analyze live sports data—making events more immersive, intelligent, and insightful.

# REFERENCES

[1] J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," *arXiv preprint arXiv:1804.02767*, Apr. 2018. [Online].

[2] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, "YOLOv4: Optimal Speed and Accuracy of Object Detection," *arXiv preprint arXiv:2004.10934*, Apr. 2020. [Online].

[3] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2015, pp. 1440–1448. [Online].

[4] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2015. [Online].

[5] C. Y. Wang, A. Bochkovskiy, and H. Y. M. Liao, "YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors," *arXiv preprint arXiv:2207.02696*, Jul. 2022. [Online].

[6] A. Cioppa, A. Deliege, S. Giancola, B. Ghanem, M. Van Droogenbroeck, R. Gade, and T. B. Moeslund, "A Context-Aware Loss Function for Action Spotting in Soccer Videos," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR) Workshops*, 2020, pp. 13126–13136.

[7] C. Y. Wang, J. S. Boon, and H. Y. M. Liao, "YOLOv9: Speed-Aware Feature Matching and Enhanced Gradient Flow for High-Efficiency Object Detection," *arXiv preprint arXiv:2402.13616*, Feb. 2024. [Online].

[8] M. Thakur, S. Singh, and K. Jain, "Real-Time Football Analytics Using YOLO and DeepSORT," *Procedia Comput. Sci.*, vol. 199, pp. 120–127, 2022.

[9] Ultralytics, "YOLOv8 Documentation and Benchmarks," 2023. [Online]. Available: https://docs.ultralytics.com/models/yolov8/

[10] Ultralytics, "YOLOv9 - Ultralytics YOLO Documentation," 2024. [Online]. Available: https://docs.ultralytics.com/models/yolov9/

[11] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2016, pp. 779–788. [Online].

[12] A. Z. Tayyebi, “Comparing YOLOv10, YOLOv9, and YOLOv8: A Performance Study,” *Medium*, 2024. [Online]. Available: https://medium.com/@tayyebi/comparing-yolov10-yolov9-and-yolov8-a-performance-study-2024

[13] S. Gino, “Sport-Analytics with YOLO et al.,” *Medium*, 2023. [Online]. Available: https://medium.com/@gino/sport-analytics-with-yolo-et-al-2023

[14] K. Moutselos and I. Maglogiannis, “Setting a Baseline for Long-Shot Real-Time Player and Ball Detection in Soccer Videos,” in *Proc. 14th Int. Conf. Inf., Intell., Syst. Appl. (IISA)*, Volos, Greece, 2023, pp. 1–7. [Online].

[15] M. Sapienza, F. Cuzzolin, and P. H. S. Torr, “Learning Object Detection in Video Surveillance,” *IEEE Trans. Circuits Syst. Video Technol.*, vol. 24, no. 3, pp. 356–371, Mar. 2014.

[16] T. Zhang, X. Zhang, Y. Yang, Z. Wang, and G. Wang, “Efficient Golf Ball Detection and Tracking Based on Convolutional Neural Networks and Kalman Filter,” *arXiv preprint arXiv:2012.09393*, Dec. 2020. [Online].

[17] D. Barry, M. Shah, M. Keijsers, H. Khan, and B. Hopman, “xYOLO: A Model for Real-Time Object Detection in Humanoid Soccer on Low-End Hardware,” *arXiv preprint arXiv:1910.03159*, Oct. 2019. [Online].

[18] M. Yaseen, “What is YOLOv9: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector,” *arXiv preprint arXiv:2409.07813*, Sep. 2024. [Online].

[19] R. Sapkota, Z. Meng, M. Churuvija, X. Du, Z. Ma, and M. Karkee, “Comprehensive Performance Evaluation of YOLO11, YOLOv10, YOLOv9, and YOLOv8 on Detecting and Counting Fruitlet in Complex Orchard Environments,” *arXiv preprint arXiv:2407.12040*, Jul. 2024. [Online].

[20] R. Sapkota, M. Pokharel, and J. Rawat, “Performance Evaluation of YOLOv8–YOLOv11 for Dense Object Detection,” *arXiv preprint arXiv:2407.12040*, Jul. 2024. [Online]. Available:

[21] K. Moutselos and I. Maglogiannis, “Setting a Baseline for Long-Shot Real-Time Player and Ball Detection in Soccer Videos,” in *Proc. 14th Int. Conf. Inf., Intell., Syst. Appl. (IISA)*, Volos, Greece, 2023, pp. 1–7. [Online]

[22] A. Bochkovskiy, C. Y. Wang, and H. Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” *arXiv preprint arXiv:2004.10934*, Apr. 2020. [Online].

# APPENDIX

## Originality Report

Report-1

ORIGINALITY REPORT

| 24% | 20% | 13% | 16% |
|---|---|---|---|
| SIMILARITY INDEX | INTERNET SOURCES | PUBLICATIONS | STUDENT PAPERS |

PRIMARY SOURCES

| | | |
|---|---|---|
| 1 | Submitted to KIET Group of Institutions, Ghaziabad<br>Student Paper | 4% |
| 2 | arxiv.org<br>Internet Source | 2% |
| 3 | www.coursehero.com<br>Internet Source | 1% |
| 4 | informatica.si<br>Internet Source | 1% |
| 5 | www.techscience.com<br>Internet Source | 1% |
| 6 | Submitted to University of Glamorgan<br>Student Paper | 1% |
| 7 | Behzad Mahaseni, Erma Rahayu Mohd Faizal, Ram Gopal Raj. "Spotting Football Events Using Two-Stream Convolutional Neural Network and Dilated Recurrent Neural Network", IEEE Access, 2021<br>Publication | <1% |
| 8 | uwaterloo.ca<br>Internet Source | <1% |
| 9 | upk.ptsb.edu.my<br>Internet Source | <1% |
| 10 | Ken D. Gorro, Elmo B. Ranolo, Anthony S. Ilano, Deofel P. Balijon. "Pothole Detection: A Study of Ensemble Learning and Decision Framework", International Journal of | <1% |

| | | |
|---|---|---|
| | Advanced Computer Science and Applications, 2025<br>Publication | |
| 11 | Submitted to Universiti Teknologi Malaysia<br>Student Paper | <1 % |
| 12 | Submitted to University of Strathclyde<br>Student Paper | <1 % |
| 13 | e-jabt.org<br>Internet Source | <1 % |
| 14 | www.mdpi.com<br>Internet Source | <1 % |
| 15 | Submitted to University of Wollongong<br>Student Paper | <1 % |
| 16 | Submitted to Obudai Egyetem<br>Student Paper | <1 % |
| 17 | Submitted to University of Birmingham<br>Student Paper | <1 % |
| 18 | Anderson Luiz dos Santos Safre, Alfonso Torres-Rua, Brent L. Black, Sierra Young. "DEEP LEARNING FRAMEWORK FOR FRUIT COUNTING AND YIELD MAPPING IN TART CHERRY USING YOLOv8 and YOLO11", Smart Agricultural Technology, 2025<br>Publication | <1 % |
| 19 | Submitted to ABES Engineering College<br>Student Paper | <1 % |
| 20 | Eman H. Alkhammash. "A Comparative Analysis of YOLOv9, YOLOv10, YOLOv11 for Smoke and Fire Detection", Fire, 2025<br>Publication | <1 % |
| 21 | Submitted to Heriot-Watt University<br>Student Paper | <1 % |
| 22 | Submitted to National Institute of Technology, MIZORAM | <1 % |

Student Paper

| | | |
|---|---|---|
| 23 | Submitted to University of York<br>Student Paper | <1 % |
| 24 | Submitted to UOW Malaysia KDU University College Sdn. Bhd<br>Student Paper | <1 % |
| 25 | bura.brunel.ac.uk<br>Internet Source | <1 % |
| 26 | Submitted to Meerut Institute of Engineering & Technology<br>Student Paper | <1 % |
| 27 | Submitted to University of Stirling<br>Student Paper | <1 % |
| 28 | export.arxiv.org<br>Internet Source | <1 % |
| 29 | infocenter.arm.com<br>Internet Source | <1 % |
| 30 | theses.lib.polyu.edu.hk<br>Internet Source | <1 % |
| 31 | Ranjan Sapkota, Manoj Karkee. "Comprehensive Performance Evaluation of YOLO11, YOLOv10, YOLOv9 and YOLOv8 on Detecting and Counting Fruitlet in Complex Orchard Environments", Qeios Ltd, 2024<br>Publication | <1 % |
| 32 | V. Sharmila, S. Kannadhasan, A. Rajiv Kannan, P. Sivakumar, V. Vennila. "Challenges in Information, Communication and Computing Technology", CRC Press, 2024<br>Publication | <1 % |
| 33 | eprints.intimal.edu.my<br>Internet Source | <1 % |
| 34 | iris.unitn.it<br>Internet Source | <1 % |

| | | | |
|---|---|---|---|
| 35 | www.cse.griet.ac.in<br>Internet Source | <1% |
| 36 | Submitted to Liverpool John Moores University<br>Student Paper | <1% |
| 37 | Submitted to VIT University<br>Student Paper | <1% |
| 38 | azmemory.azlibrary.gov<br>Internet Source | <1% |
| 39 | ir.uew.edu.gh:8080<br>Internet Source | <1% |
| 40 | Submitted to American University in Dubai<br>Student Paper | <1% |
| 41 | Lu-hao He, Yong-zhang Zhou, Lei Liu, Yu-qing Zhang, Jian-hua Ma. "Research on the directional bounding box algorithm of YOLO11 in tailings pond identification", Measurement, 2025<br>Publication | <1% |
| 42 | Submitted to University of Essex<br>Student Paper | <1% |
| 43 | dspace.cvut.cz<br>Internet Source | <1% |
| 44 | www.theseus.fi<br>Internet Source | <1% |
| 45 | Submitted to Hogeschool Utrecht - Tii<br>Student Paper | <1% |
| 46 | Submitted to King's College<br>Student Paper | <1% |
| 47 | Submitted to Malta College of Arts,Science and Technology<br>Student Paper | <1% |
| 48 | Submitted to The University of Manchester | |

| | | |
|---|---|---|
| | Student Paper | <1 % |
| 49 | Submitted to University of Bristol<br>Student Paper | <1 % |
| 50 | dspace.daffodilvarsity.edu.bd:8080<br>Internet Source | <1 % |
| 51 | intellipaat.com<br>Internet Source | <1 % |
| 52 | sear.unisq.edu.au<br>Internet Source | <1 % |
| 53 | Correia, Tales Veríssimo Souza. "Detection and Mosaicing Through Deep Learning Models for Low-Quality Retinal Images", Instituto Politecnico de Leiria (Portugal), 2024<br>Publication | <1 % |
| 54 | Yi Tianjing, Mustafa Mhamed. "Developments in Automated Harvesting Equipment for the Apple in the Orchard: Review", Smart Agricultural Technology, 2024<br>Publication | <1 % |
| 55 | digitalcommons.usf.edu<br>Internet Source | <1 % |
| 56 | etd.uum.edu.my<br>Internet Source | <1 % |
| 57 | sist.sathyabama.ac.in<br>Internet Source | <1 % |
| 58 | www.1.ise-iq.org<br>Internet Source | <1 % |
| 59 | www.awfarlak.com<br>Internet Source | <1 % |
| 60 | bitspace.co.in<br>Internet Source | <1 % |
| 61 | doczz.net<br>Internet Source | |

<1 %

| | | |
|---|---|---|
| 62 | repositorio.uloyola.es<br>Internet Source | <1 % |
| 63 | www.inderscience.com<br>Internet Source | <1 % |
| 64 | www.nature.com<br>Internet Source | <1 % |
| 65 | www.scirp.org<br>Internet Source | <1 % |
| 66 | 9pdf.net<br>Internet Source | <1 % |
| 67 | Arvind Dagur, Karan Singh, Pawan Singh Mehra, Dhirendra Kumar Shukla. "Intelligent Computing and Communication Techniques - Volume 1", CRC Press, 2025<br>Publication | <1 % |
| 68 | Brij B. Gupta, Michael Sheng. "Machine Learning for Computer and Cyber Security - Principles, Algorithms, and Practices", CRC Press, 2019<br>Publication | <1 % |
| 69 | Gaurav Aggarwal, Ashutosh Tripathi, Himani Goyal Sharma, Tripti Sharma, Rishabh Dev Shukla. "Integrated Technologies in Electrical, Electronics and Biotechnology Engineering", CRC Press, 2025<br>Publication | <1 % |
| 70 | Zainab Fatima, M. Hassan Tanveer, Hira Mariam, Razvan Cristian Voicu, Tanazzah Rehman, Rizwan Riaz. "Performance Comparison of Object Detection Models for Road Sign Detection Under Different Conditions", International Journal of | <1 % |

| | | | |
|---|---|---|---|
| | Advanced Computer Science and Applications, 2024<br>Publication | |
| 71 | assets-eu.researchsquare.com<br>Internet Source | <1 % |
| 72 | cronfa.swan.ac.uk<br>Internet Source | <1 % |
| 73 | ntnuopen.ntnu.no<br>Internet Source | <1 % |
| 74 | www2.mdpi.com<br>Internet Source | <1 % |
| 75 | zealpress.com<br>Internet Source | <1 % |
| 76 | Antonella Santone, Francesco Mercaldo, Luca Brunese. "A Method for Real-Time Lung Nodule Instance Segmentation Using Deep Learning", Life, 2024<br>Publication | <1 % |
| 77 | Submitted to Brunel University<br>Student Paper | <1 % |
| 78 | Carlos Guedes Soares, Tiago A. Santos. "Advances in Maritime Technology and Engineering - Volume 1", CRC Press, 2024<br>Publication | <1 % |
| 79 | Santos, Tomás Moreira. "Real-Time Weapon Detection in Surveillance Footage", Universidade do Porto (Portugal), 2023<br>Publication | <1 % |
| 80 | Tang, Zhicheng. "New Deep Learning Methods for Small to Medium Object Detection in Images", University of Missouri - Columbia, 2024<br>Publication | <1 % |
| 81 | core.ac.uk<br>Internet Source | <1 % |

| | | |
|---|---|---|
| 82 | gitarattan.edu.in<br>Internet Source | <1 % |
| 83 | idoc.pub<br>Internet Source | <1 % |
| 84 | oresta.org<br>Internet Source | <1 % |
| 85 | repository.uwtsd.ac.uk<br>Internet Source | <1 % |
| 86 | repositum.tuwien.at<br>Internet Source | <1 % |
| 87 | tudr.thapar.edu:8080<br>Internet Source | <1 % |
| 88 | www.frontiersin.org<br>Internet Source | <1 % |
| 89 | www.jetir.org<br>Internet Source | <1 % |
| 90 | www.kluniversity.in<br>Internet Source | <1 % |
| 91 | Vibhav Sachan, Shahid Malik, Ruchita Gautam, Kumar Parvin. "Advances in AI for Biomedical Instrumentation, Electronics and Computing", CRC Press, 2024<br>Publication | <1 % |
| 92 | Vinod M. Kapse, Lalit Garg, Pavan Kumar Shukla, Varadraj Gurupur, Amit Krishna Dwivedi. "Applications of Artificial Intelligence in 5G and Internet of Things", CRC Press, 2025<br>Publication | <1 % |
| 93 | scholarsmine.mst.edu<br>Internet Source | <1 % |
| 94 | Burhan Duman. "A Real-Time Green and Lightweight Model for Detection of Liquefied | <1 % |

Petroleum Gas Cylinder Surface Defects Based on YOLOv5", Applied Sciences, 2025
Publication

| | | |
|---|---|---|
| 95 | Hoang-Tu Vo, Kheo Chau Mui, Nhon Nguyen Thien, Phuc Pham Tien. "Automating Tomato Ripeness Classification and Counting with YOLOv9", International Journal of Advanced Computer Science and Applications, 2024<br>Publication | <1 % |
| 96 | dx.doi.org<br>Internet Source | <1 % |
| 97 | hdl.handle.net<br>Internet Source | <1 % |
| 98 | journal.uad.ac.id<br>Internet Source | <1 % |
| 99 | scholarworks.uark.edu<br>Internet Source | <1 % |

| | | | |
|---|---|---|---|
| Exclude quotes | Off | Exclude matches | Off |
| Exclude bibliography | Off | | |

# Proof of Submission

**5th International Conference on Computational Methods in Science & Technology 2025 : Submission (209) has been created.**

2 messages

**Microsoft CMT** <noreply@msr-cmt.org> Mon, May 5, 2025 at 6:41 PM
To: aditya.2125cse1127@kiet.edu

Hello,

The following submission has been created.

Track Name: ICCMST2025

Paper ID: 209

Paper Title: Comparative Analysis of YOLOv8 and YOLOv9 for Real-Time Object Detection in Live Football Broadcasts

Abstract:
This study contrasts the two advanced models, YOLOv8 and YOLOv9, for object detection, in the demanding area of real-time object identification during live football broadcasts. We assessed their performance using a tailored dataset of football images, concentrating on detecting players, the ball, and the boundary of the field. The research analyses the key attributes of each detection model, including detection accuracy, processing speed, and computing efficiency, while evaluating the advantages and disadvantages of each model. These findings will elucidate its appropriateness in sports analytics and further the evolution of real-time object detection in live sports settings.

Created on: Mon, 05 May 2025 13:10:59 GMT

Last Modified: Mon, 05 May 2025 13:10:59 GMT

Authors:
- aditya.2125cse1127@kiet.edu (Primary)

Primary Subject Area: Track 3 - Machine Learning

Secondary Subject Areas: Not Entered

Submission Files:
Comparative Analysis of YOLOv8 and YOLOv9 for Real-Time Object Detection in Live Football Broadcasts.docx (144 Kb, Mon, 05 May 2025 13:10:43 GMT)
Comparative Analysis of YOLOv8 and YOLOv9 for Real-Time Object Detection in Live Football Broadcasts.pdf (237 Kb, Mon, 05 May 2025 13:10:55 GMT)

Submission Questions Response: Not Entered

Thanks,
CMT team.

**Research Paper**

# Comparative Analysis of YOLOv8 and YOLOv9 for Real-Time Object Detection in Live Football Broadcasts

Eva Mittal*, Devyanshi Srivastava†, Aditya Agnihotri‡ and Gaurav Parashar§
Department of Computer Science and Engineering, KIET Group of Institutions
Ghaziabad, India
Email: *evamittal2003@gmail.com, †devyanshi.2125cse@kiet.edu, ‡aditya.2125cse1127@kiet.edu, §gauravparashar24@gmail.com

***Abstract*—This study contrasts the two advanced models, YOLOv8 and YOLOv9, for object detection, in the demanding area of real-time object identification during live football broadcasts. We assessed their performance using a tailored dataset of football images, concentrating on detecting players, the ball, and the boundary of the field. The research analyzes the key attributes of each detection model, including detection accuracy, processing speed, and computing efficiency, while evaluating the advantages and disadvantages of each model. These findings will elucidate its appropriateness in sports analytics and further the evolution of real-time object detection in live sports settings.**

## I. Introduction

Real-time object identification has transformed the realm of sports analytics, enabling automated tracking, event detection, and performance assessment with greater precision and speed. This technology has proven crucial for applications such as player tracking, ball identification, and tactical evaluation, considerably enhancing decision making in both coaching and broadcasting [1], [2]. From the many object recognition frameworks, deep learning models as a family have demonstrated great efficiency, with the You Only Look Once (YOLO) family emerging as the preferred choice as it offers the best of both worlds in terms of speed as well as accuracy. [3]. Later iterations of YOLO have consistently improved detection precision, feature extraction, and processing efficiency, making them especially suitable for fast-paced environments such as live football broadcasts [4], [5].

Prior studies have thoroughly examined the earlier iterations of YOLO, including YOLOv3 and YOLOv4, in the context of sports, highlighting their effectiveness in identifying players and analyzing games [6], [7]. However, limited research have studied the latest iterations, YOLOv8 and YOLOv9, in the setting of live football events. These versions bring major architectural advancements, including transformer-based feature extraction and redesigned detection pipelines, which promise higher accuracy and real-time performance [8], [9].

This study conducts a comparative evaluation of YOLOv8 and YOLOv9 for real-time object detection in live football broadcasts. By examining crucial performance measures, including detection accuracy, inference speed, and computing economy, we seek to evaluate the relative advantages of each model in high-speed, complex visual environments. Our findings contribute to improve real time sports analytics by identifying the top effective deep learning models for automated event detection and tactical assessment in football.

## II. Related Work

Object detection has progressed markedly through deep learning models, notably contributing to sports analytics by facilitating real-time tracking and automatic broadcasting. The YOLO series has been thoroughly examined for its efficacy at identifying multiple items in dynamic settings.

Earlier YOLO versions have been applied in sports analytics. [1] proposed YOLO as a real-time object detection system, with YOLOv3 [2] increasing feature extraction. [3] demonstrated YOLOv4's efficacy in football player tracking, leveraging CSPDarknet-53 for enhanced detection under diverse conditions. [4] showed YOLOv5's efficiency in multi-object tracking with reduced computational overhead.

Despite improvements, real-time football analytics still faces difficulties such as occlusions, fast player movements, and fluctuating camera angles. Studies by [5] and [6] demonstrated that early YOLO systems failed with fast moving objects, underlining the necessity for better feature extraction and transformer-based architectures. Recent YOLO versions introduce major enhancements. [7] demonstrated YOLOv8's superiority in real-time sports applications, offering improved FPS and accuracy. [8] highlighted YOLOv9's robustness against motion blur and occlusion, making it ideal for football analytics. YOLOv9's adaptive attention mechanisms enhance player tracking and ball trajectory estimation.

Although YOLOv8 and YOLOv9 represent major improvements in the YOLO series with improved architectures, enhanced feature extraction capabilities, and optimized detection accuracy, little is known about how well they perform against one another for demanding tasks like event detection during football games and live player tracking. To understand how effectively these state-of-the-art models perform in dynamic, real-time sporting environments, this gap must be closed.

## III. Methodology

### A. Dataset

In order to compile a comprehensive and representative dataset for our study, we carefully collected a variety of images obtained from live football broadcasts. This dataset aimed to encompass a diverse selection of real-life scenarios seen during football matches, ensuring that the object detection models were trained on highly dynamic and diverse visual inputs.

The dataset features dynamic action sequences that effectively portray the intensity and quick movements of players in a live game. These sequences document intricate interactions like dribbling, passing, tackling, and attempting goals, which are essential for training models to identify detailed player movements and ball trajectories.

Additionally, wide-angle perspectives were embedded to provide a comprehensive understanding of the game, allowing the models to learn team formations, player positioning, and overall field dynamics. These wide views mimic real broadcast angles, which improve the model's effectiveness during full match analysis.

To achieve the most granular level of detection, we included close-up images, which focus on discrete players, referees and critical moments in the game. This allowed the model to detect subtle features like player expressions, jersey numbers, and particular actions like kicking, or diving.

For precise model training and evaluation, every image in the dataset was meticulously analyzed. Bounding boxes were carefully drawn around each detected object: players, ball, and field boundaries, ensuring that the models received essential spatial information for accurate object localization and recognition.

These enhanced qualitative annotations improve the detection capabilities of YOLOv8 and YOLOv9 by providing accurate training instances of the complexities of real-life football broadcasts. The expansive dataset is vital for testing the performance of YOLOv8 and YOLOv9 on live sports analysis, ensuring that the models are ready to deal with the challenges of real football matches.

### B. Models

To enhance the training efficiency and fully utilize the data at our disposal, we utilized pretrained versions of YOLOv8 and YOLOv9 as the foundation for our object detection frameworks. These pretrained models have been trained on extensive and diverse image datasets, equipping them with a profound understanding of visual characteristics, including edges, textures, object shapes, and spatial relationships. This comprehensive prior training allows the models to identify patterns and structures prevalent in many images, rendering them highly efficient in object detection tasks.

To customize these models for our specific football object recognition task, we employed transfer learning techniques, which enable the adaptation of existing knowledge to a new domain while minimizing computing expenses and training duration. Transfer learning utilizes feature representations that have already been learned from pre-trained models and adjusts them with our specific dataset, allowing the models to become proficient at identifying objects pertinent to live football broadcasts.

The optimization process involved retraining the models in our curated dataset, which consists of various scenarios from football matches, including player movements, ball tracking, and field boundary recognition. By modifying the model parameters during the fine-tuning stage, we improved their capability to accurately identify and pinpoint objects within the football setting, even when confronted with rapid movements, changing lighting conditions, and intricate player interactions. This strategic combination of pre-trained models and domain-specific fine-tuning resulted in a more efficient and high-performing object detection system. Using transfer learning, we ensured that our models not only benefited from the robustness of prior knowledge but also seamlessly adapted to the intricacies of real-time football analytics, making them well suited for deployment in live sports broadcasts.

## IV. Evaluation Metrices

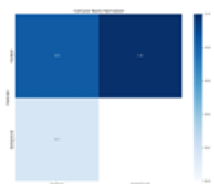
In this part, we discuss on the assessment criteria and visualizations that would be used to evaluate the performance of YOLOv8 and YOLOv9. These measures are critical to assessing the strengths and shortcomings of the models in detecting objects during live football matches. In the following, we provide thorough explanations for the selected charts and their significance.

### A. Confusion Matrix (Normalized)

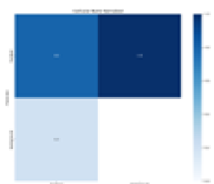
The normalized confusion matrix provides insights into the performance of the model by presenting the proportions of properly and erroneously identified occurrences of objects. Normalization stresses relative performance over absolute measurements, enabling comparison between YOLOv8 and YOLOv9.

Significance:

- Highlights the true positive, false positive, true negative, and false negative rates.
- Offers clarity on model precision and recall across various object categories.



(a) YOLOv8 Confusion Matrix



(b) YOLOv9 Confusion Matrix

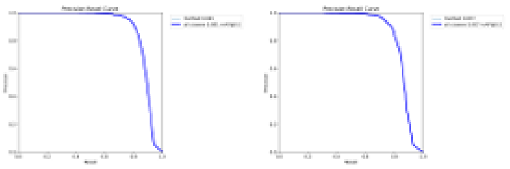
Fig. 1: Comparison of confusion matrices for YOLOv8 and YOLOv9

### B. *Precision-Recall (PR) Curve*

The PR curve offers an extensive perspective on the trade-off between precision and recall at various levels. It is particularly advantageous for imbalanced datasets where elevated recall may result in lower precision.

Significance:

- Demonstrates the ability of models to balance precision and recall.
- Indicates thresholds where the models perform optimally.



(a) YOLOv8 PR Curve (b) YOLOv9 PR Curve

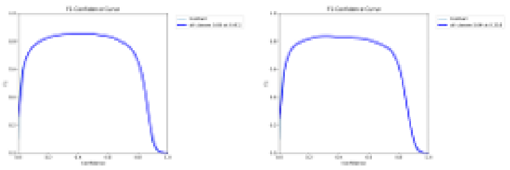
Fig. 2: Comparison of PR Curve for YOLOv8 and YOLOv9

### C. *F1 Confidence Curve*

The F1 confidence curve provides a reasonable assessment of a model's performance by combining precision and recall into a single measure. This curve helps discover the appropriate thresholds to maximize performance.

Significance:

- Highlights the harmonic mean of precision and recall.
- Provides a clear metric for comparing YOLOv8 and YOLOv9.



(a) YOLOv8 F1 Curve (b) YOLOv9 F1 Curve

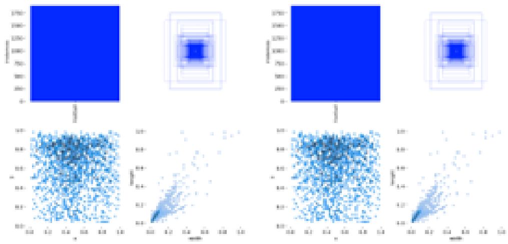
Fig. 3: Comparison of F1 Curve for YOLOv8 and YOLOv9

### D. *Labels Distribution*

The Labels Distribution chart displays the frequency of object types within the dataset. This information is critical for comprehending dataset bias and its potential impact on model performance.

Significance:

- Identifies underrepresented or overrepresented categories.
- Helps assess the alignment of dataset composition with real-world scenarios.



(a) YOLOv8 Labels (b) YOLOv9 Labels

Fig. 4: Comparison of Labels for YOLOv8 and YOLOv9

## Results

### *Confusion Matrix*

The classification performance of YOLOv8 and YOLOv9 across categories such as "football", "player" and "background" is depicted in normalized confusion matrices (Fig. 1).

- YOLOv8: Achieved greater true positive rates, with 548 correct classifications for "football" and fewer misclassifications compared to YOLOv9.
- YOLOv9: Performed similarly but demonstrated slightly greater misclassifications, with 540 accurate classifications for "football" and a substantially poorer precision.

### *Precision-Recall (PR) Curve*

The PR curve (Fig. 2) indicates the trade-off between precision and recall at varying confidence thresholds.

- YOLOv8: Demonstrates a superior performance with an mAP@0.5 of 0.881, showing robust precision and recall across all classes.
- YOLOv9: Achieves an mAP@0.5 of 0.857, slightly lower than YOLOv8, indicating minor inconsistencies in balancing precision and recall.

### *F1 Confidence Curve*

The F1 confidence curve (Fig. 3) evaluates the balance between precision and recall at varying thresholds.

- YOLOv8: Achieves an F1 score of 0.86 at a confidence threshold of 0.411, thus demonstrating superior balance and performance stability.
- YOLOv9: Peaks at an F1 score of 0.84 at a confidence threshold of 0.318, indicating a trade-off favoring recall at lower thresholds.

### *Labels Distribution*

The placement and density of object labels throughout the dataset is depicted in the label distribution comparison (Fig. 4) between YOLOv8 and YOLOv9. With dense label clusters in central area of the frames, both models demonstrate a comparable distribution pattern. This suggests that during live broadcasts, the majority of objects of interest, such as football players and the ball, are concentrated in central areas.

- YOLOv8: Labels are equally dispersed across the image with fewer outliers, therefore demonstrating consistent detection across the dataset.
- YOLOv9: Label placement shows slight irregularities compared to YOLOv8, with a marginally higher spread in edge cases.

## Discussion

### *Confusion Matrix*

YOLOv8 demonstrates slightly better classification accuracy than YOLOv9, with fewer false positives and false negatives. This indicates that YOLOv8 is better at correctly identifying objects within the dataset, making it more reliable for tasks where precision is critical.

### *Precision-Recall (PR) Curve*

The greater mAP@0.5 of YOLOv8 demonstrates its superior ability to balance precision and recall. This is critical in real-time applications where false positives or false negatives can dramatically impair decision-making, such as player tracking or ball detection.

### *F1 Confidence Curve*

The better F1 score and appropriate confidence threshold of YOLOv8 imply a more balanced performance. YOLOv9, while significantly less balanced, may still perform effectively in cases where recall is favored above precision, such as tracking moving players across numerous frames.

### *Labels Distribution*

The label distribution research demonstrates that both YOLOv8 and YOLOv9 are well-suited for football analytics, as their label placement patterns correspond with the spatial distribution of essential features in live football broadcasts. YOLOv8's regularity in label placement suggests higher stability in detecting things of interest.

## Conclusion

This study conducted a thorough evaluation of YOLOv8 and YOLOv9 for object detection in live football broadcasts, analyzing key performance metrics such as mean Average Precision (mAP), precision-recall curves, F1-scores, and label distribution. Our findings reveal that while both models demonstrate strong capabilities in real-time sports analytics, YOLOv8 exhibits a slight overall edge in performance.

Specifically, YOLOv8 delivers higher classification accuracy, striking a more optimal balance between precision and recall. This balance is crucial in live sports scenarios, where real time decision making and minimal false detections are paramount. Furthermore, YOLOv8 maintains a more consistent label distribution across various test cases, suggesting greater robustness in identifying players, the ball, and field boundaries, even in dynamic and challenging match conditions. This stability enhances its reliability for practical applications such as player tracking, ball movement analysis, and automated highlight generation.

On the other hand, YOLOv9 demonstrates promising results, particularly in scenarios where recall is a priority. Its ability to detect a higher number of true positives makes it a viable option for applications where capturing every possible object instance is more critical than minimizing false positives. However, when considering the overall balance of speed, accuracy, and computational efficiency, YOLOv8 emerges as the more attractive choice for real-world deployment in sports analysis. These insights contribute to the ongoing advancement of real time object detection in sports analytics, offering valuable guidance for selecting the most suitable model based on specific application requirements. Future research could explore hybrid approaches that integrate the strengths of both YOLO versions to further optimize performance in live sports detection tasks.

## Future Work

### *Real-time Implementation*

Assessing the real-time capabilities of YOLOv8 and YOLOv9 on high-resolution video feeds should be the aim of future studies. This includes tracking latency, processing rates per frame, and inference speed during live operations. The goal is to ascertain if these models can operate at the necessary speed for live football broadcasts without encountering significant delays or computational limitations while ensuring high accuracy.

### *Multi-object Tracking*

For improving the effectiveness of these models in sports analytics, it is essential to combine multi-object tracking (MOT) algorithms with YOLOv8 and YOLOv9. By employing tracking techniques such as DeepSORT or ByteTrack, future research can improve the ability to consistently track multiple players and the ball across continuous frames. This will enable more advanced applications, including trajectory forecasting, movement heatmaps, and analysis of team strategies.

### *Robustness Analysis*

An important topic that needs further study is how resilient these models are in different situations. This means assessing their performance under different lighting conditions (e.g., day versus night games, stadium floodlights versus natural light), different camera angles (e.g., close-ups, sideline shots, and aerial views), and player attire variations (e.g., different team uniforms or referee interference). Gaining an understanding of these elements will enhance YOLO models' adaptability in practical applications.

### *Dataset Expansion*

To enhance model generalization and performance, expanding the dataset to include a broader range of football scenarios is crucial. The dataset should encompass a variety of weather conditions (including rain, fog, and snow), different camera perspectives (such as broadcast cameras, drone shots, and cameras worn by players), and unique playing styles (like fast-paced attacking versus defensive setups). A larger dataset will improve the accuracy of the model and yield trustworthy results in a range of real-world scenarios.

By addressing these areas, future research can further optimize YOLO-based object detection systems for real-time football analytics, making them more efficient, accurate, and applicable to professional sports technology solutions.

## References


[1] Redmon, J., Divvala, S., Girshick, R., and Farhadi, A. (2016). "You Only Look Once: Unified, Real-Time Object Detection."

[2] Ultralytics. (2024). "YOLOv9 - Ultralytics YOLO Documentation."

[3] Tayyebi, A.Z. (2024). "Comparing YOLOv10, YOLOv9, and YOLOv8: A Performance Study." Medium.

[4] Gino, S. (2023). "Sport-Analytics with YOLO et al." Medium.

[5] Moutselos, K., and Maglogiannis, I. (2023). "Setting a Baseline for Long-Shot Real-Time Player and Ball Detection in Soccer Videos."

[6] Sapienza, M., Cuzzolin, F., and Torr, P.H.S. (2014). "Learning Object Detection in Video Surveillance." IEEE Transactions on Circuits and Systems for Video Technology, 24(3), 356-371.

[7] Zhang, T., Zhang, X., Yang, Y., Wang, Z., and Wang, G. (2020). "Efficient Golf Ball Detection and Tracking Based on Convolutional Neural Networks and Kalman Filter." arXiv preprint arXiv:2012.09393.

[8] Ultralytics. (2023). "YOLOv8 - Ultralytics YOLO Documentation." This document offers comprehensive information on the YOLOv8 model, including its design and applications.

[9] Barry, D., Shah, M., Keijsers, M., Khan, H., and Hopman, B. (2019). "xYOLO: A Model For Real-Time Object Detection In Humanoid Soccer On Low-End Hardware." arXiv preprint arXiv:1910.03159. This paper introduces xYOLO, an adaptation of the YOLO model designed for real-time object detection in humanoid soccer robots, emphasizing performance on low-end hardware.

[10] Yaseen, M. (2024). "What is YOLOv9: An In-Depth Exploration of the Internal Features of the Next-Generation Object Detector." arXiv preprint arXiv:2409.07813. This paper offers a comprehensive analysis of the YOLOv9 model, detailing its architectural innovations, training methodologies, and performance improvements over previous versions.

[11] Sapkota, R., Meng, Z., Churuvija, M., Du, X., Ma, Z., and Karkee, M. (2024). "Comprehensive Performance Evaluation of YOLO11, YOLOv10, YOLOv9, and YOLOv8 on Detecting and Counting Fruitlet in Complex Orchard Environments." arXiv preprint arXiv:2407.12040. This study evaluates various YOLO versions, including YOLOv8 and YOLOv9, for detecting and counting fruitlets in complex orchard environments, providing insights into their performance.